

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number
WO 02/20814 A1

(51) International Patent Classification: **C12N 15/86**,
15/861, 5/10, 15/11, 15/63, 15/64, 15/65, A61K 48/00

Eun-Chung [KR/US]; 91 Waltham Street #1, Boston, MA
02118 (US). WANG, Xinzhong [CN/US]; 4 Woodmere
Road, Framingham, MA 01701 (US).

(21) International Application Number: PCT/US01/27682

(74) Agent: **ELBING, Karen, L.**; Clark & Elbing LLP, 176
Federal Street, Boston, MA 02110-2214 (US).

(22) International Filing Date:
7 September 2001 (07.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/231,053 8 September 2000 (08.09.2000) US
60/246,904 8 November 2000 (08.11.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:

US 60/231,053 (CON)
Filed on 8 September 2000 (08.09.2000)
US 60/246,904 (CON)
Filed on 8 November 2000 (08.11.2000)

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

(71) Applicant (*for all designated States except US*): **THE
GENERAL HOSPITAL CORPORATION** [US/US]; 55
Fruit Street, Boston, MA 02214 (US).

Published:

- with international search report
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SEED, Brian**
[US/US]; 9 Hawthorne Place #5J, Boston, MA 02114
(US). **WRIGHT FREEMAN, Mason** [US/US]; 203
Lincoln Road, Lincoln, MA 01773 (US). **KOVTUN,
Alexander** [US/US]; 156 Arlington Street, Acton, MA
01720 (US). **MURAKAWA, Masahiro** [JP/US]; 506
Beacon Street, Apt. 6, Boston, MA 02115 (US). **PARK,**

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: SELF-REARRANGING DNA VECTORS

(57) Abstract: Disclosed are replicatable viral DNA vectors encoding a site-specific DNA-altering enzyme and a DNA target recog-
nized by the enzyme, the enzyme selectively converting, in a cell expressing the enzyme, the DNA vector to a rearranged form. The
invention further relates to methods for assembling recombinant adenoviral DNAs. These methods include the steps of: providing
a first linearized DNA vector including a restriction site and a cos site and a second linearized DNA vector including the restriction
site, an adenoviral nucleic molecule, and a cos site; and ligating the first and second linearized DNA vectors, the ligation assembling
a recombinant adenoviral DNA.

WO 02/20814 A1

5

SELF-REARRANGING DNA VECTORS

Background of the Invention

The invention relates to DNA vectors.

10 Mammalian cell expression vectors based on DNA viruses have been widely discussed as gene delivery vehicles for genetic therapy. Among the different DNA viruses proposed for this purpose have been adenoviruses, baculovirus, Epstein Barr virus, and herpes simplex virus. In addition other smaller viruses that have an intranuclear phase in which the viral genome is present as a double stranded DNA, such as retroviruses and parvoviruses, have been proposed as gene delivery vehicles.

Adenoviral vectors (AdV), for example, have a recognized potential for gene delivery, founded in their broad host range, robust growth in culture, and capacity to infect mitotically quiescent cells (Graham and Prevec, Manipulation of adenovirus vectors, p. 109-128, *In* E. J. Murray (ed.), Methods in Molecular Biology, vol. 7, Humana, Clifton, NJ, 1991; Trapnell and Gorziglia, Curr. Opin. Biotechnol. 5:617-625, 1994). AdV can be propagated in a helper cell line, 293, a human embryonic kidney cell line transformed by adenovirus type 5 (Graham et al., J. Gen. Virol. 36:59-72, 1994). 293 cells express the viral E1 gene products (E1a and E1b) that are the master regulatory proteins for subsequent viral gene expression. E1 deleted viruses can propagate in 293 cells, but not in other cells. Although it would be expected that E1 deleted viruses lack the machinery to express viral genes, several studies have demonstrated that cellular E1-like components can stimulate viral gene expression (Imperiale et al., Mol. Cell. Biol. 4:867-74, 1984; Onclercq et al., J. Virol. 62:4533-7, 1988; Spergel et al., J. Virol. 66:1021-30, 1992). The expression of these viral genes results in the relatively rapid elimination of transduced cells *in vivo* as a result of cytotoxic T cell responses (Yang et al., Immunity 1:433-42, 1994; Yang et al., Gene Ther. 3:137-44, 1996; Yang et al., J. Virol. 69:2004-15, 1995).

Thus attention has been focused on eliminating the remaining vestiges of viral expression. Viral genes that have been deleted for this purpose include the gene for E4 proteins (Armentano et al., Hum. Gene Ther. 6:1343-53, 1995; Kochanek et al., Proc.

Natl. Acad. Sci. USA 93:5731-6, 1996; and Yeh et al., J. Virol. 70:559-565, 1996), DNA binding protein (Engelhardt et al., Proc. Natl. Acad. Sci. USA 21:6196-6200, 1994; and Gorziglia et al., J. Virol. 70:4173-8, 1996), DNA polymerase (Amalfitano et al., J. Virol. 72:926-33, 1998), and the preterminal protein (Schaack et al., Proc. Natl. Acad. Sci. USA 93:14686-91, 1996). The most aggressive approach has been the creation of helper virus-dependent vectors that lack all viral genes (Hardy et al., J. Virol. 71:1842-9, 1997; Kochanek et al., Proc. Natl. Acad. Sci. USA 93:5731-6, 1996; Lieber et al., J. Virol. 70:8944-60, 1996; Mitani et al., Proc. Natl. Acad. Sci. USA 92:3854-8, 1995; and Parks et al., Proc. Natl. Acad. Sci. USA 93:13565-13570, 1996). These vectors have high capacity, evoke reduced cellular immune responses and show prolonged expression *in vivo* (Morsy et al., Proc. Natl. Acad. Sci. USA 95:7866-71, 1998). However to deploy these viruses on the scale required for human clinical application presents major challenges because a cesium chloride (CsCl) gradient is needed to remove the helper virus.

15

Summary of the Invention

In one aspect, the invention features a replicatable viral DNA vector encoding a site-specific DNA-altering enzyme and a DNA target recognized by said enzyme, said enzyme selectively converting, in a cell expressing said enzyme, said DNA vector to a rearranged form.

20

In preferred embodiments, the rearranged form includes an autonomously replicating episome and a linear DNA product. In other preferred embodiments, the vector comprises adenoviral DNA.

In yet other preferred embodiments, the vector includes a genetically-engineered recombination site (such as a target of Cre or FLP). Preferably, such a recombination site includes a recognition sequence of a site-specific DNA altering enzyme.

25

In another preferred embodiment, the site-specific DNA altering enzyme is a recombinase (such as Cre or FLP) or an integrase. Preferably, such an enzyme is functional in a mammalian cell. Preferred embodiments of the vector also include an origin of replication that functions in a mammalian cell (such as an Epstein Barr Virus replicon). Moreover, the vector typically includes a gene of interest (such as a therapeutic gene that encodes a protein or polypeptide or an RNA product).

30

In another aspect, the invention features a method for assembling a recombinant adenoviral DNA. The method, in general, includes the steps of: (a) providing a first linearized DNA vector comprising a restriction site and a cos site and a second linearized DNA vector comprising the restriction site, an adenoviral nucleic acid molecule, and a cos site; and (b) ligating the first and second linearized DNA vectors, the ligation assembling a recombinant adenoviral DNA.

In preferred embodiments, the first linearized DNA vector comprises a selectable marker (such as a gene encoding a polypeptide that confers, on a host cell expressing such a polypeptide, resistance to an antibiotic). In other preferred embodiments, the first linearized DNA vector includes an adenoviral left-end inverted terminal repeat; a gene of interest; or both. In still other preferred embodiments, the second linearized DNA vector includes a selectable marker. Preferably, the second linearized DNA vector includes an adenoviral right-end inverted terminal repeat.

The method further includes packaging the assembled adenoviral DNA into a phage and infecting a host cell. Typically the first and second linearized DNAs include cosmid vector DNA. In addition, such adenoviral DNA is typically flanked by cleavage sites (such as intron endonuclease cleavage sites).

In another aspect, the invention features an adenovirus producer cell having a nucleic acid molecule that expresses a dominant negative site-specific DNA-altering enzyme. In preferred embodiments, the site-specific DNA altering enzyme is a dominant negative recombinase (for example, a Cre recombinase such as CreY324C or a Flp recombinase). Exemplary adenovirus producer cells include, without limitation, 293 human embryonic kidney cells, per.C6 cells, and N52 cells.

In yet another aspect, the invention features a vector comprising, in the 5' to 3' direction, a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; a gene of interest; a lineage-specific gene promoter; a second genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; and a nucleic acid molecule encoding a site-specific DNA altering enzyme.

In still another aspect, the invention features a vector including, in the 5' to 3' direction, a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; a gene of interest; a bi-directional promoter, comprising a second genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; and a nucleic acid molecule encoding a site-specific DNA altering enzyme.

In related aspects, the invention features a method of gene therapy including the administration to a patient in need of gene therapy a therapeutically effective amount of the vector of the invention, which is expressed in the patient. The invention further relates to a population of cells transfected with the vector of the invention.

5 Accordingly, the invention further relates to the use of a recombinant viral vector or use of a recombinant viral particle for gene therapy. Such vectors and viral particles may be introduced either *in vitro* into a host cell removed from the patient, or directly *in vivo*, into the body to be treated, according to standard methods known in the art.

10 The invention also relates to a pharmaceutical composition that includes a therapeutically effective amount of a recombinant viral vector or viral particle prepared according to the methods disclosed herein, in combination with a vehicle that is acceptable from a pharmaceutical standpoint. Such a pharmaceutical composition may be prepared according to the techniques commonly employed and administered by any known administration route, for example systemically (in particular, by intravenous,
15 intratracheal, intraperitoneal, intramuscular, subcutaneous, intratumoral, or intracranial routes) or by aerosolization or intrapulmonary administration.

 One skilled in the art will appreciate that suitable methods of administering a vector (particularly an adenoviral vector) of the present invention to an animal for purposes of gene therapy, chemotherapy, and vaccination are available, and, although
20 more than one route can be used for administration, one particular route may provide a more immediate and more effective reaction than another. Pharmaceutically acceptable excipients also are well known to those who are skilled in the art, and are readily available. The choice of excipient will be determined, in part, by the particular method used to administer the recombinant vector or particle. Accordingly, there are a wide
25 variety of suitable formulations for use in the context of the present invention.

 By "recombinant DNA vector" is meant a DNA sequence containing a desired sequence (such as a gene of interest) and an appropriate regulatory element(s) necessary for the expression of the operably linked sequence in a particular host organism (such as a mammal).

30 By "operably linked" is meant that a gene and a regulatory element(s) are connected to permit gene expression when the appropriate molecules (for example, transcriptional activator proteins) are bound to the regulatory sequence(s).

By "regulatory element" is meant a genetic element that controls some aspect of the expression of a nucleic acid sequence. For example, a promoter is a regulatory element that facilitates the initiation of transcription of an operably linked coding region. Other genetic regulatory elements include, without limitation, splicing signals, polyadenylation signals, and termination signals. For example, transcriptional regulatory elements in eukaryotes include promoter and enhancer elements. Promoters and enhancers include arrays of DNA sequences that interact directly or indirectly with cellular proteins involved in transcription. Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in mammalian cells and viruses.

By "transfection" is meant the introduction of foreign DNA into eukaryotic cells. Transfection is typically accomplished by a variety of means known in the art including, without limitation, calcium phosphate-DNA co-precipitation, DEAE-dextran-mediated transfection, electroporation, microinjection, liposome fusion, lipofection, protoplast fusion, and biolistics.

By "stably transfected" is meant the introduction of foreign DNA into the genome of the transfected cell. In general, transfer and expression of transgenes in mammalian cells are now routine practices to those skilled in the art, and have become major tools to carry out gene expression studies and to generate vectors useful in gene therapy.

By "gene of interest" is meant a gene inserted into a vector whose expression is desired in a host cell. Genes of interest include, without limitation, genes having therapeutic value, as well as reporter genes. A variety of such genes are useful in the invention, including genes of interest encoding a protein, which provides a therapeutic function. In addition, the gene of interest, if a therapeutic gene, can render its effect at the level of RNA, for instance, by encoding an antisense message or ribozyme, a protein which affects splicing or 3' processing (e.g., polyadenylation), or it can encode a protein which acts by affecting the level of expression of another gene within the cell (i.e., where gene expression is broadly considered to include all steps from initiation of transcription through production of a processed protein), for example, by mediating an altered rate of mRNA accumulation, an alteration of mRNA transport, and/or a change in post-transcriptional regulation.

By "reporter gene" is meant a gene sequence that encodes a reporter molecule (including an enzyme). A "reporter molecule" is detectable in any detection system, including, but not limited to, enzyme (e.g., ELISA, as well as enzyme-based histochemical assays), fluorescent, radioactive, and luminescent systems. Exemplary
5 reporter gene systems include the *E. coli* beta-galactosidase or glucuronidase genes, green fluorescent protein (GFP), blue fluorescent protein (BFP), the human placental alkaline phosphatase gene, the chloramphenicol acetyltransferase (CAT) gene; other reporter genes are known in the art and may be employed as desired.

By "transgene" is meant any piece of DNA, which is inserted by artifice into a
10 cell, and becomes part of the genome of the organism, which develops from that cell. Such a transgene may include a gene that is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an endogenous gene of the organism.

By "transgenic" is meant any cell that includes a DNA sequence, which is
15 inserted by artifice into a cell and becomes part of the genome of the organism, which develops from that cell.

By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation).

By "derived from" is meant isolated from or having the sequence of a naturally
20 occurring sequence (e.g., a cDNA, genomic DNA, synthetic, or combination thereof).

By "nucleic acid" is meant a polynucleotide (DNA or RNA).

By "gene" is meant any nucleic acid sequence coding for a protein or an RNA molecule.

By "gene product" is meant either an untranslated RNA molecule transcribed
25 from a given gene or coding sequence (such as, mRNA or antisense RNA) or the polypeptide chain translated from the mRNA molecule transcribed from the given gene or coding sequence. Nucleic acids according to the invention can be wholly or partially synthetically made, can comprise genomic or complementary DNA (cDNA) sequences, or can be provided in the form of either DNA or RNA.

30 The presently claimed invention affords a number of advantages. For example, applicants' gene therapy vehicles particularly those based on recombinant adenoviruses, minimize the propensity of the vectors to activate host immune surveillance, and thereby maximize the persistence for the DNA transduced. The invention therefore facilitates

the development of gene delivery vectors designed to enhance persistence of virally delivered genes and evade the cellular immune response by severing the connection between the sole adenoviral enhancer and the sequences encoding potentially antigenic viral proteins.

- 5 As described in more detail below, the mechanism by which this is accomplished differs significantly from any other previous approaches. For example, to reduce the immunogenicity of vectors it is widely acknowledged that some intervention, such as the removal of key genes, or the prevention of their expression in the cells targeted for therapy, is important; however, many related approaches are directed at the host and
- 10 have generally focused on the selective induction of tolerance to adenoviral antigens, or similar strategies directed at inducing a temporally restricted or antigen-specific compromise of the immune system.

- In addition, the poor persistence of transduced DNA appears to be due in part to immunological rejection of transduced cells and to the inability of the viral DNA to
- 15 replicate, a feature generally inherent in the design of adenoviral vectors, but one which is not associated with applicants' claimed gene therapy vehicles.

- Moreover, some contemporary adenoviral vectors are designed to propagate in specific host cells which provide essential replication factors *in trans*. These vectors are typically based on cell lines which express the master regulatory proteins of the E1
- 20 complex, which are required for induction of adenoviral DNA replication. In cells expressing E1 genes, the best studied of which is a human embryonic kidney cell line transformed by DNA from human adenovirus 5 (called HEK293, or simply 293), viruses lacking E1 genes propagate well. Such viruses do not propagate on cell lines which do not express E1, and do not generally propagate well in the target cells to which the
- 25 therapeutic gene is to be delivered. Cells transduced with E1-deleted adenovirus vectors also do not express high levels of viral genes in the absence of E1. However, the weak residual expression that remains in such vectors appears to be sufficient to induce cellular immune responses that contribute to the destruction of the transduced cells.

- In addition, the gene therapy vectors claimed herein are hybrid vectors capable of
- 30 self-rearrangement to form circular and linear DNA products. The linear DNA has a compromised ability to express adenoviral genes, and therefore has a lower immunological profile. And the circular DNA behaves like a mammalian plasmid, encoding the gene of interest and persisting by autonomous replication in the nucleus.

For example, the circularization of an adenoviral vector via the action of Cre recombinase beneficially places a gene of interest (for example, a therapeutic gene) on a self-replicating episome. Vector circularization occurs in a tissue-targeted manner, for example, as a result of the activation of a synthetic liver-specific promoter upstream of the recombinase Cre. Once circularized, the EBV replicon in the episome confers improved persistence on the therapeutic gene as detected by reporter gene expression and direct assay for the presence of vector DNA sequences.

Furthermore, the invention eliminates the requirement for a helper virus, thus avoiding two potential limitations of that system. First, the continuous expression of Cre recombinase may lead to toxicity in host cells, either as a direct consequence of the protein's activity or via its immunogenicity. Second, the Cre helper virus may itself produce antigenic viral proteins that contribute to the immunologic elimination of infected host cells. In contrast, the self-resolving adenovirus/EBV vector system disclosed herein advantageously provides no alternative source of viral proteins, and Cre expression is terminated upon rearrangement.

In addition, the invention described herein provides tools for analyzing the roles of the enhancer in viral gene regulation and virus growth.

The invention also provides a convenient general system for creating recombinant adenoviruses, which increase their attractiveness as gene transduction tools for basic research. The system, for example, employs two conventional plasmid vectors and a λ phage packaging step. The entire recombinant AdV genome is assembled into a single cosmid that is easily amplified in *E. coli*. The use of intron endonuclease recognition sequences flanking the ITRs enhances virus production while simplifying insertion of therapeutic gene sequences into the pLEP shuttle plasmid. The convenience of this vector system has facilitated the construction of over two hundred recombinant viruses to date.

Other embodiments and advantages of the invention will be apparent from the detailed description thereof, and from the claims.

Brief Description of the Drawings

FIGURE 1A is a schematic diagram of the structure of an adenoviral type A vector and its fate in a target cell. enh refers to the Ad2 enhancer; GFP refers to the marker gene green fluorescent protein; EBV refers to the Epstein Barr Virus replicon;

TetO₇ refers to a heptamer of Tet operator; TetR refers to the Tet repressor; VP16 refers to the viral protein 16 of Herpes simplex virus, SD refers to the splice donor site; and SA refers to the splice acceptor site.

FIGURE 1B is a schematic diagram of the structure of an adenoviral type B vector and its fate in a target cell. enh refers to the Ad2 enhancer; GFP refers to the marker gene green fluorescent protein; EBV refers to the Epstein Barr Virus replicon; SD refers to the splice donor site; and SA refers to the splice acceptor site.

FIGURE 2A shows a schematic diagram of the pLEP cosmid polylinker region and its position relative to the adenoviral left ITR. The adenovirus enhancer/packaging sequence (ψ) is boxed.

FIGURE 2B is a schematic diagram showing the generation of a single cosmid encoding the AdV genome by the direct ligation of two smaller plasmids. A gene expression unit, CMVGFP, was inserted into the pLEP cosmid at the polylinker region. pLEP and pREP cosmids were digested with an intron endonuclease (PI-PspI), ligated, and packaged *in vitro* to generate pAd2CMVGFP. This DNA was then digested with another intron endonuclease (I-CeuI) to expose the ITRs at both ends of the viral genome. Finally, cosmid digestion mixtures were transfected into 293 cells. Plaques generated by recombinant viruses are detected in 7-10 days.

FIGURE 3A shows the restriction analysis of cosmids carrying the full length AdV DNA showing uniform generation of the desired vector DNA. 2 μ g DNA samples from four pAd2-7CMVGFP colonies were digested with Bgl II, resolved on a 1% agarose gel and stained with ethidium bromide. The predicted sizes of the DNA fragments are: 13261, 7684, 5228, 5088, 2284, 1757, 1549, 1270, 351, and 275 base pairs (bp). The 5228 and 5088 fragments appear as a doublet, and the 351 and 275 bp fragments are too small to be seen on the gel.

FIGURE 3B shows the release of the recombinant Ad DNA from cosmids by I-CeuI digestion. 2 μ g of pAd2-7CMV DNA from two clones was digested with I-CeuI. Arrows indicated the position of the released recombinant AdV DNA and the vector fragments of approximately 35 kb and 5 kb, respectively.

FIGURE 4A shows the appearance of plaques in 293 cells transfected with 10 μ g of pAdGFPB with no ITR exposed (undigested), one ITR exposed (BsaBI or I-CeuI), or both ITRs exposed (BsaBI plus I-CeuI). Values represent the mean plaque counts per

dish and the time required for plaque development in 293 cells from three separate experiments. "T" designates I-CeuI; and "B" designates Bsa BI.

FIGURE 4B shows the viral titers obtained from plaques that were allowed to grow over 10 days after transfection. Viruses were harvested and the titer of each virus stock was determined by a GFP based semi-quantitative titration procedure described herein. Values represent the mean \pm SE of three independent determinations.

FIGURE 5 is a schematic diagram showing a linear AdV that resolves into a circular episome. The elements involved in the self-directed rearrangement of the vector are shown schematically in pLEP1BHCRGFP/EBV and in the corresponding AdV.

Starting from the left ITR, the elements are shown as following sequence: left ITR, 147 bp; first 34 bp loxP site; 185 bp enhancer/packaging signal; 64 bp splicing acceptor (SA) from EF1 α gene first intron; 720 bp GFP cDNA; 230 bp SV40 poly(A); 1.7 kb TK-EBNA-1/OriP; 970 bp HCR12 promoter; 1 kb EF1 α gene first intron containing splicing donor (SD) and acceptor (SA) sites with the second loxP site inserted at 64 bp upstream of the 3' end; 1.2 kb Cre gene tagged with AU1 and a nuclear localization signal; ~120bp poly(A) signal and PI-PspI site. After infection of liver cells, the HCR12 promoter drives the expression of Cre which results in the cleavage of the two loxP sites. This results in the circularization of the fragment containing the EBV replicon. The excision severs the connection between the enhancer/packaging signals and the remainder of the AdV genome. The Cre gene becomes promoterless and is left on the AdV genome fragment. After excision, the HCR12 promoter drives the expression of the GFP reporter gene. The EBV replicon maintains the excised circle as an episome in host cells.

FIGURE 6A is a schematic representation of the loxP sites and EBNA-1 locations in the AdV genome. The relevant Bgl II site is also shown.

FIGURE 6B shows the time course of rearrangement in HepG2 and Hela cells at an equal multiplicity of infection (moi) of 1,000 particles per cell. Cells were infected with Ad2HCRGFP/EBV viruses for 2 hours at 37 °C. Hirt DNA samples were extracted from the cells. ~5 μ g of Hirt DNA samples were digested with Bgl II, fractionated on a 1% agarose gel, and analyzed by Southern blot techniques using a ³²P-labeled EBNA-1 fragment as the hybridization probe.

FIGURE 6C shows the DNA blot results obtained from HeLa cells infected at a moi of 10,000; and HepG2 at 1,000. The upper bands (4915 bp) represent the circularized DNA fragments whereas the lower bands (3162 bp) represent the non-circularized AdV.

5 FIGURE 7A shows green fluorescent protein (GFP) expression in liver and non-liver cells infected with the Ad2HCRGFP/EBV viruses. Cells were cultured in 35 mm dishes and infected with the Ad2HCRGFP/EBV virus at desired moi. HepG2 cells were infected with 1,000 particles per cell, whereas HeLa cells were infected with a moi of 10,000. GFP expression was examined at the indicated time points after infection.

10 Fluorescent cells were photographed using an Olympus SC35mm camera mounted on an Olympus IX70 fluorescent microscope, at 200x magnification, using a filter with peak excitation and emission wavelengths of 450 nm and 510 nm, respectively.

FIGURE 7B shows the expression of GFP in HepG2, HeLa, A431, and HT29 cells. Cells were seeded in 35 mm dishes and infected with the Ad2HCRGFP/EBV virus at a moi of 10,000 particles per cell. GFP expression was examined at 72 hours after infection.

15

FIGURE 7C shows the expression of GFP in human primary hepatocytes. These cells were photographed under bright field (left) and fluorescent conditions (right).

FIGURE 8A shows the results of RT-PCR that was performed to detect the tripartite leader sequence (upper panel) for virus late gene expression; and PCR was performed in the DNA samples for detection of the AdV genomes. The specific target sequences are described in detail *infra*. PCR analyses of adenovirus late gene expression in cells infected with the first generation AdVs or the self-resolving Ad2HCRGFP/EBV was analyzed. HepG2 cells were cultured in 35 mm dishes and infected with increasing moi (0, 10, 100, 1000, 10,000, and 100,000) of adenoviral vectors. RNA and DNA were isolated in parallel from the cells at 72 hours after infection.

20

25

FIGURE 8B shows a summary of quantitative RT-PCR and PCR results. Each determinant was the average of three experiments.

FIGURE 9A is a schematic diagram depicting the deletion analysis of the OriP and EBNA-1 regions of the EBV replicon. Structures of the deletions in EBNA-1 and OriP are schematically represented. Elements considered important for episomal maintenance are indicated. FR refers to the family of repeats; DS designates the region

30

of dyad symmetry; LR1 refers to the so-called linker region 1; GA refers to gly-ala repeats; LR2 refers to linker region 2; and Dimerization designates the dimerization domain.

FIGURE 9B is a graph depicting fractions of GFP positive cells carrying the
5 EBV replicons represented in Figure 9A.

FIGURE 10A shows the positions and identities of Cre mutants tested for their dominant negative Cre activities.

FIGURE 10B is a schematic diagram of the substrate Cre plasmid (ad2239) used to test dominant negative functions of Cre mutants.

10 FIGURE 10C shows GFP expression in cells cotransfected with the substrate Cre plasmid (ad2239) and the indicated Cre mutants.

FIGURES 10D and 10E show Cre mutants tested for their ability to inhibit rearrangement. Only those showing the strongest inhibitory activities were retested in Fig. 10E. GFP intensity was normalized to that of cells in the absence of inhibition.

15 FIGURE 11 shows GFP expression in 293 TetON cells and #17 cells transfected with ad2239. The ability of #17 cells to inhibit Cre activity is demonstrated by the weak GFP signal in cells treated with 2 μ M doxycycline.

FIGURE 12 is a schematic diagram depicting the tetracycline mediated auto-regulatory circuit.

20 FIGURES 13A and 13B show the effects of different basal elements on synthetic TetO promoter activity. FIGURE 13A shows a schematic diagram of the components of various auto-regulatory synthetic TetO promoters. FIGURE 13B shows a comparison of the strength of auto-regulatory synthetic TetO promoters bearing different basal elements, in the presence and absence of tetracycline, using GFP as a marker in HepG2
25 cells.

FIGURE 14 shows the structure of a Cre substrate plasmid (ad2265). The promoter, Efl α , and the gene, BFP, are interrupted by two loxP sites, which can be joined by Cre-mediated recombination. PA stands for poly A; BFP for blue fluorescent protein.

30 FIGURES 15A and 15B show the estrogen regulation of Cre recombinase activity. 293 cells infected with type B virus, AD121.5, in which the Cre enzyme is fused with estrogen ligand binding domain at both the N- and C-termini were cultured in the presence or absence of 1 μ M estrogen. Cre-mediated rearrangement in the presence

of estrogen is shown in Figure 15A, whereas blot analysis of extrachromosomal DNA from the same cells is shown in Figure 15B. L represents the position corresponding to the unarranged adenoviral DNA; and C represents the position corresponding to the circular form of DNA.

5 FIGURE 16 shows the rearrangement of adenoviral sequences *in vivo*. Extrachromosomal DNA from the livers of Rag-2 mice sacrificed 2.5 hrs post injection of type A adenovirus, AD102.7, was analyzed by DNA blot. L represents the size corresponding to linear adenoviral DNA; and C represents the size corresponding to rearranged circular DNA.

10 FIGURE 17 is a photomicrograph depicting high level GFP expression in Rag2 mouse hepatic tissues 48 hrs post type A adenovirus (AD102.7) injection.

 FIGURE 18A and 18B show schematic diagrams of the structures of adenoviral vectors and their fates in target cells. enh refers to Ad2 enhancer; GFP refers to green fluorescent protein; EBV refers to Epstein Barr Virus replicon; TetO₇ refers to heptamer
15 of Tet operator; TetR refers to Tet repressor; VP16 refers to transcriptional activator domain from HSV protein 16; SD refers to splice donor site; and SA refers to splice acceptor site.

 FIGURE 19A shows the structure of a FLP substrate plasmid, ad2879. The promoter, Ef1 α , and the gene, GFP, are interrupted by 2 FRT sites, which can be joined
20 by the FLP-mediated recombination. PA stands for poly A; BFP for blue fluorescent protein.

 FIGURE 19B shows the structure of a cre substrate plasmid, ad2204.

 FIGURE 20 shows the structures of several FLPe anti-sense plasmids.

 FIGURE 21 is a panel of photomicrographs showing inhibition of FLP enzyme
25 activity by anti-sense FLP. 293 cells were transfected with FLP substrate (Figure 12) and plasmids indicated in each photo. High GFP intensity indicate the higher expression of FLP and less inhibition by the anti-sense expressed.

 FIGURE 22 shows a schematic diagram of FRT/Cre and loxP/FLP adenovirus.

30

Detailed Description

Described herein are systems for the regulated self-rearrangement of DNA vectors, for example, gene therapy vectors. Such regulated self-rearrangement has the potential to prevent unwanted expression of vector genes not required for a therapeutic effect, and to allow the stable association of the therapeutic gene with the target cell.

The essential elements of the regulated DNA rearrangement system are a gene which encodes one or more proteins which induce DNA rearrangement, a method for regulating the activity of those proteins or their abundance, and a target DNA sequence on which those proteins act. Particularly desirable are methods for regulating the activity of the proteins or their abundance which can be easily carried out on an intact organism, such as administration or withdrawal of a drug, hormone, or environmental stimulus such as heat or irradiation, which induces the activity or abundance of the proteins which cause DNA rearrangement.

Especially desirable are regulated DNA rearrangement systems in which all of the components can be delivered in a single vector. An example of this is a virus which bears both the cis-acting sequences for DNA rearrangement as well as the protein or proteins which act on those sequences, and the regulatory apparatus which controls the activity or abundance of those proteins. However, it is not necessary that the different elements be encoded in a single nucleic acid.

The important elements of this strategy are: the compromise of vector gene function by regulated rearrangement of DNA topology, the generation of plasmid circles from vector DNA in a regulated manner, and the removal of enhancer or promoter elements from the vector DNA by regulated excision. It is also important that the circular DNA generated by site-specific recombination possesses a mechanism for stable association with the host genome in some form, here conferred by the EBV replicon. In other embodiments, the circular DNA might possess the ability to direct its integration into the host chromosomes by a site-specific integration. Site-specific integration into the host chromosomes may also be generated by the action of a regulated site-specific recombinase on a linear template without passing through a circular intermediate.

Also described herein is one particular self-rearranging vector that begins as a hybrid adenovirus vector which is capable of converting itself into two unlinked molecules, a circular and a linear DNA. After this event the linear DNA product is deleted for two important cis-acting sequences: the packaging signals, which are

required for insertion of the viral DNA into the viral capsid, and the enhancer, which increases the expression of other promoters encoded in the viral DNA. The remaining linear DNA is thereby compromised in its ability to express adenoviral genes, endowing the vector with a lower immunological profile. The circular DNA generated by the
5 excision event is a mammalian cell plasmid which has the capacity to persist by autonomous replication in the nucleus. This capacity is encoded in genetic elements derived from the Epstein Barr virus (EBV). A schematic diagram of such a vector is illustrated in Figure 5.

Epstein Barr virus is a human herpes virus which is the etiologic agent of
10 infectious mononucleosis and which has been implicated in the genesis of Burkitt's lymphoma, a B cell neoplasm, and is thought to be a predisposing factor for some forms of nasopharyngeal carcinoma. Approximately 85% of the adult Western population has a persistent population of B cells which contain a circular latent form of the viral-genome, maintained in cells by the action of Epstein Barr Nuclear Antigen 1 (EBNA1),
15 a DNA replication protein that acts on the viral latent phase origin of replication, OriP. EBNA1 in and of itself is not thought to promote neoplasia; current thinking places greater weight on the actions of the EBNA2 proteins and LMP, latent membrane protein, in the inception of EBV-associated neoplasm.

Mammalian cell plasmids have been created which bear the EBNA1 gene and
20 OriP. In nonrodent cells, these plasmids persist by replication with each transit of the cell cycle. Multiple transcription units can be borne by these plasmids, allowing regulated expression of diverse gene products.

Preferred adenoviral vectors, shown in Figures 1A and 1B, are linear forms of an EBV plasmid flanked by loxP sites, cis-acting sequences required for site-specific
25 recombination directed by the bacteriophage P1 cre protein. To prepare an adenovirus bearing both the cre protein and loxP sites, it is necessary to insure that the cre protein is not expressed while the vector is being propagated in 293 cells. To lower the immunological profile of the vector, it is also desirable that the cre protein not be expressed after the vector delivered its payload to the target cell and the cre protein
30 performed its function.

To accomplish these objectives, two general approaches have been developed for the production of adenoviral chromosomes that circularize following the regulated expression of site-specific recombinases. In each case, the vector is engineered to allow

for the production of viruses in 293 cells, and to provide transitory expression of recombinase that induces rearrangement in target tissues. The major difference between the two strategies lies in the way the deinduction of recombinase is achieved.

In the first approach, adenoviral vectors are engineered to turn an activating
5 transcription factor into a repressor upon chromosomal rearrangement. Vectors
employing this approach are referred to herein as type A vectors (Figure 1A). In the
second approach, the recombinase promoter is redirected following chromosomal
rearrangement. Vectors utilizing the second approach are referred to as type B vectors
(Figure 1B). In both cases a linear chromosome is converted to its circular episomal
10 form and a resulting deleted linear form. The circular DNA contains an Epstein Barr
virus (EBV) replicon, which allows synchronous replication of the episome with the host
mitotic cycle (Reisman et al., Mol. Cell Biol. 8: 1822-32, 1985; Yates et al., Nature 313:
812-15, 1985). The linear DNA is deleted for the enhancer and E1 genes.

One self-regulated gene switch, employing the type A vector strategy, was
15 designed based on the bacterial transposon Tn10 tetracycline repressor (tetR) gene. In
its natural context, the tetR protein binds to specific sequences (tet operator sequences)
upstream of a tetracycline resistance gene, preventing transcription of the gene unless
tetracycline is present. To adapt this protein for eukaryotic gene regulation, a gene
fusion is created between tetR and an active portion of a strong eukaryotic
20 transcriptional activator, the herpes simplex virus VP16 protein. The fusion protein
exerts its action on a synthetic promoter created by the insertion of multiple tet operator
sequences upstream of a basal promoter element. This configuration allows high-level
gene expression whenever the tetR-VP16 fusion protein binds to its cognate operator
sequences. Because the tetR protein normally does not bind to its operator in the
25 presence of tetracycline, the activity of this synthetic promoter is high in the absence of
tetracycline and low in its presence.

One example of a type A vector is shown in Figure 1A. This self-regulated gene
expression cassette, present in a hybrid adenovirus, consists of a bi-directional promoter
element in which central tetR binding sites are flanked by divergently oriented basal
30 promoter elements. In one direction the promoter directs the formation of a transcript
encoding the cre protein; in the other direction, the promoter directs the formation of a
tetR-VP16 fusion protein. The latter differs from the conventional version in bearing a
loxP site between the tetR component and the VP16 component. When tetracycline is

present this gene switch is silent. As shown in Figure 1A, upon introduction into a target cell in the absence of tetracycline, the tetR-loxP-VP16 fusion protein is produced, stimulating further production of the fusion protein, and the cre protein. The cre protein then acts to promote site specific recombination between the loxP site in the tetR-loxP-
5 VP16 coding sequences, and a distant loxP site. As a result of this recombination, the fusion protein coding sequence is disrupted so that the promoter no longer directs the formation of a tetR-loxP-VP16 fusion protein, but gives rise to an inert tetR-loxP-VP16 fusion protein for binding to the promoter upstream elements, thereby extinguishing promoter activity.

10 As shown in Figure 1A, the excised circular DNA element contains at least two transcription units. In addition, other transcription units or internal ribosome entry site elements may be used to allow the coexpression of gene products which are useful for extending the persistence of the delivered DNA, regulating expression of the gene of interest, or providing for ablation of the transduced cells once their presence is no longer
15 desirable. In addition, the linear DNA remaining after excision of the circular gene expression plasmid lacks both viral packaging sequences and the cis-acting enhancer. Within this linear DNA, additional loxP sites may be placed to provide for the rearrangement of the remaining vector DNA in the target cell, disrupting the normal topology of the genes, and further thwarting expression.

20 Using the type B vector design strategy, described in greater detail below, a recombinant adenoviral gene delivery system that is capable of undergoing growth phase-dependent site-specific recombination has also been constructed.

The following examples are presented for the purpose of illustrating, not limiting, the invention.

25

TYPE B VECTORS – EXPERIMENTAL RESULTS

Several experimental examples for constructing type B vectors and for carrying out the general approaches of the invention are now described below.

30 Two-Cosmid System for Efficient Construction of Recombinant AdV

To simplify and facilitate the generation of recombinant AdV, a system was established to assemble the desired AdV genome in a single plasmid by ligation (shown in Figures 2A, 2B). The system consists of two component vectors, a left end plasmid,

pLEP, and a right end plasmid, pREP. The left end Ad sequences (nt 1-376) in pLEP include the viral inverted terminal repeat, the cis-acting packaging sequences, and the viral enhancer. Nucleotide (nt) positions described herein refer to the wild type Ad2 sequence in GenBank (J019017). The Ad sequences are followed by the gene
5 expression unit intended for delivery and an intron endonuclease (PI-PspI) cleavage site. The right end plasmid contains a PI-PspI site followed by the Ad2 genome from the end of the E1 locus rightward (nt 3527-35937).

pLEP is a small tractable vector for cloning, whereas pREP is much larger and contains less frequently manipulated genes. Both pLEP and pREP contain a
10 bacteriophage λ *cos* site, oriented to generate a single cosmid of appropriate length for *in vitro* packaging following ligation of the two plasmids at the PI-PspI cleavage site. pLEP is tetracycline resistant (Tet^r) and pREP is ampicillin (Amp^r) resistant, allowing the recombinants to be selectively isolated by co-selection for both markers. In the resulting assembled cosmid, the adenoviral sequences are closely flanked by cleavage
15 sites for the intron endonuclease I-CeuI. Digestion with I-CeuI liberates the entire recombinant AdV genome from the parent cosmid (see Figure 2B).

Three classes of pREP have been constructed to allow the preparation of AdVs bearing E1 (pREP7; SEQ ID NO.: 2), E1 and E3 (pREP8; SEQ ID NO.: 3), or E1, E3, and E4 (pREP12; SEQ ID NO.: 4) deletions. pREP7 (SEQ ID NO.: 2) contains nt
20 3527-35937 of the Ad2 genome, and pREP8 (SEQ ID NO.: 3) carries an additional deletion in the E3 region (Δ nt 27901-30841). pREP12 (SEQ ID NO.: 4) has deleted open reading frames (ORF) 1-4 of the E4 region (Δ nt 34121-35469, 1348 bp). AdV generated with these cosmids should be able to accommodate 5, 8, and 10 kb inserts, respectively.

25 These aforementioned vectors were constructed as follows. The EcoRI to BsaI fragment that spans the ampicillin resistance gene in pBR322 was deleted and replaced by a synthetic adapter, and the bacteriophage λ *cos* site was inserted between the unique StyI and BsmI sites. A PCR amplified Ad2 fragment containing the left end ITR (L-ITR), enhancer elements, and the encapsidation signal (nt 1-376) was created and
30 inserted into the adapter (Figures 2A, 2B) to yield the tetracycline-resistant left-end plasmid pLEP. The right end of Ad2 from the AflII site to the right end (nt 3527-35937) was assembled into an ampicillin resistant cosmid vector, pACKrr3 (SEQ ID NO.: 1), by multiple steps of PCR amplification and fragment interchange. The resultant cosmid

was termed pREP7 (SEQ ID NO.: 2). To expand vector capacity, two deletions were incorporated into the pREP7 (SEQ ID NO.: 2) cosmid, an E3 gene deletion (nt 27901-30841, 2840 bp); cosmid pREP8 (SEQ ID NO.: 3) and a 1.3 kb deletion (nt 34121-35469) in the E4 region of the Ad2 region; pREP12 (SEQ ID NO.: 4)

5 An example of the construction of an AdV carrying a CMV-GFP expression unit is outlined in Figure 2. pLEPCMVGFP (Tet^r) was digested with P1-PspI and ligated to the pREP7 (SEQ ID NO.: 2; ΔE1, Amp^r) digested with the same enzyme. The ligation mixture was packaged with λ phage extracts (MaxPlax lambda packaging extracts, Epicentre Technologies) and a fraction of the packaged phage was used to infect a
10 recombination-deficient *E. coli* host, with selection for the assembled plasmid on Amp/Tet plates. Transductants containing pLEP fused to pREP were selected on agar containing 25 μg/ml ampicillin and 12.5 μg/ml tetracycline (Amp/Tet). Colonies were selected and DNA isolated (Qiagen). DNA was used either for restriction analysis or for tranfection of 293 cells as described herein.

15 Figure 3A shows typical results for the Bgl II digestion pattern of a pLEP3CMVGFP/pREP7 hybrid cosmid, pAd2-7CMVGFP DNA. Because of the size minimum (~40 kbp) for λ phage *in vitro* packaging and the double antibiotic selection, most of the colonies growing on Amp/Tet plates were the desired hybrid cosmids, and undesired rearrangements were rarely seen. In the present example, all four pAd2-
20 7CMVGFP clones exhibited the digestion pattern predicted from the inferred sequence. The entire recombinant AdV genome was then released from the cosmid by I-CeuI digestion (Figure 3B). I-CeuI digestion leaves ten nucleotides to the left of the left ITR and eight nucleotides to the right of the right ITR. Short flanking sequences have been reported to be eliminated during replication of recombinant viruses after transfecting the
25 DNA into 293 (human embryonic kidney) cells (Hanahan et al., Mol. Cell. Biol. 4:302-309, 1984).

 The digestion reaction can be transfected into 293 cells without purification as follows. 293 cells, obtained from Microbix Biosystems (Ontario, Canada), were cultured in 10 cm dishes in complete Dulbecco's Modified Eagle's Medium (DMEM)
30 supplemented with 10% FBS, 2mM glutamine and penicillin/streptomycin (Gibco BRL), and maintained at 37 °C and 5% CO₂ atmosphere incubator. Cells were grown to ~50% confluence on the day of transfection. Ten μg of cosmid DNA were digested with I-CeuI in a volume of 50 μl. The reaction mixture was transfected into 293 cells by

calcium phosphate precipitation (Graham and Prevec, Manipulation of adenovirus vectors, p. 109-128, *In* E. J. Murray (ed.), Methods in Molecular Biology, vol. 7, Humana, Clifton, NJ, 1991) without purification. After transfection, cells were cultured and examined daily for the appearance of cytopathic effects (CPE). Virus propagation, purification, plaque assay, and viral DNA isolation were performed using established protocols (Graham and Prevec, *supra*). At day six post-transfection, 5-30 viral plaques/10 cm dish /10 μ g DNA were usually apparent, which compared favorably with the 30-50 plaques/10 cm dish /10 μ g DNA found for 293 cells transfected with purified wild type Ad2 DNA.

10 To compare the efficiency of recombinant virus production, similar viruses were also generated by homologous recombination. 20 μ g of pREP7 (SEQ ID NO.: 2) was co-transfected into 293 cells with 10 μ g of a plasmid encoding the left end of the adenoviral genome and a green fluorescent reporter gene (pLITREF1 α GFP). pLITREF1 α GFP contained the Ad2 left end nt 1-376, an EF1 α promoter/GFP expression unit and Ad2 sequence (from 3525-8120) that overlaps with the same sequence in pREP7 (SEQ ID NO.: 2). This overlap fragment served as the region for homologous recombination. Each co-transfection was performed in duplicate. Initial plaques took longer to appear (14 days post transfection) and were less abundant (0-3 plaques per plate).

20 Data in the literature suggest that exposed ITR ends favor efficient virus production (Hanahan et al., *supra*). To assess the importance of this effect, an AdV cosmid, pIADEF1 α GFPB, in which the AdV ITRs were flanked with a different restriction site at each end was constructed. pIADEF1 α GFPB DNA was digested with BsaBI to expose the right ITR, I-CeuI to expose the left ITR, or the two enzymes were used together to expose both ends. Digested cosmid DNA samples were transfected into 293 cells and plaques were allowed to develop. Virus propagation, purification, plaque assay, and viral DNA isolation were performed using established protocols described in Graham and Prevec. (Manipulation of adenovirus vectors, *In* E. J. Murray (ed.), Methods in Molecular Biology, vol. 7. Humana, Clifton, NJ., pp. 109-128, 1991).

30 Ten days after transfection the viruses were harvested and viral titers were determined. The average titer for the viral stocks (Figures 4A, 4B) was 1.3×10^4 pfu/ml from transfection with undigested DNA; 2.4×10^5 pfu/ml from BsaBI linearized DNA (free right ITR); 1.1×10^5 pfu/ml from I-CeuI linearized DNA (free left ITR); and $2.7 \times$

10⁶ pfu/ml for the BsaBI/I-CeuI double digested DNA (both ITRs free). Thus liberation of each end resulted in an approximate increase in the efficiency of generating virus by a factor of ten (Figures 4A, 4B).

5 Construction of an AdV Capable of Self-Rearrangement

One approach to attenuating adenoviral gene expression and improving transgene persistence is the creation of viruses capable of undergoing internal, self-directed rearrangement upon delivery to the target tissue. In principle, this objective can be achieved through the regulated expression of site-specific recombinases in vectors that
10 contain the cis-acting target of recombinase action. To allow such vectors to be created, the recombinase activity must be suppressed during propagation in the packaging cell line. As described in more detail below, the use of a lineage-specific promoter to control recombinase expression has been successfully employed to achieve this end.

An example of this is shown in Figure 5. The expression of Cre recombinase
15 was controlled by a liver-specific promoter constructed as follows. The human hepatic control region 1 and 2 (HCR1 and 2) of the ApoE/C gene locus (Allan et al., J. Biol. Chem. 270:26278-81, 1995; and Dang et al., J. Biol. Chem. 270:22577-85, 1995) were amplified by PCR using 293 cell genomic DNA as the template. The following primers were used to amplify both HCR1 and HCR2 fragment: HCRtop-
20 5'gcggaattcggcttggtgacttagagaacagag 3' (SEQ ID NO.:5); HCRbot - 5' gcgggatccttgaacccggaccctctcacacta 3' (SEQ ID NO.:6). The amplified PCR fragments (~0.39 kb) were cloned into pUC19. The HCR1 and HCR2 sequences were confirmed by dideoxy DNA sequencing. The two fragments were assembled in a head to tail orientation, fused with a synthetic basal TATA element and cloned in a parental pLEP
25 vector containing a GFP reporter gene. The resultant plasmid was named pLEPHCR12GFP. The synthetic liver-specific, as demonstrated below, provided a means to control Cre recombinase expression during propagation of the vector in 293 cells, and allowed for testing the consequences of abstracting the enhancer from the linear vector DNA upon delivery of the DNA to the target cells.

30 In 293 cells, this promoter is silent, allowing the viral chromosome to be propagated with minimal rearrangement. Any rearranged viruses that are formed lack packaging signals and so disappear from the pool of propagating vectors. In liver cells the Cre recombinase is induced by the action of the tissue-specific promoter. The

resulting Cre-induced recombination excises a circular episome and redirects the transcriptional output of the liver-specific promoter so that it directs the synthesis of the transgene of interest. The remaining linear fragment consists of an adenoviral genome lacking the enhancer and packaging signals and a Cre expression unit devoid of promoter sequences.

In the form discussed here, one loxP site is located at nucleotide 147 of the Ad2 genome, between the left ITR and the enhancer/packaging sequences, and the second loxP site is placed inside an intron a few bases upstream of the splice acceptor sequence. Hence the loxP site does not appear in the resulting mature transcript. The Cre coding sequence that remains on the right end linear fragment after rearrangement lies downstream from a splice acceptor that lacks a splice donor or upstream promoter sequences. This effectively terminates the expression of Cre following excision.

Prior to recombination, the Cre recombinase gene is under the control of a synthetic promoter (referred to as HCR12), consisting of hepatic locus control elements from the human ApoE/C locus fused to the first intron of the human EF1 α gene. After cyclization the HCR12 promoter lies upstream of the transgene (in this case GFP) and the distal segment of the intron (beyond the loxP site) contains the adenoviral enhancer. To facilitate manipulation of the plasmids in *E. coli*, the human IgG1 hinge-CH2 intron (118 bp) was inserted in the Cre coding sequence at nucleotide 237, suppressing Cre expression in bacteria. The circularized episome contains the latent origin of replication (OriP) and trans-acting DNA replication protein (EBNA-1) of Epstein Barr virus, and hence is capable of autonomous replication in synchrony with the host mitotic cycle (Yates et al., Nature 313:812-815, 1985).

Using the two cosmid system described above, the pLEP plasmid containing the self-resolving components, pLEP1BHCR12, was ligated with pREP8 (SEQ ID NO.: 3; Δ E1 Δ E3) to create pAdVHCRGFP/EBV. The latter was digested with I-CeuI and transfected into 293 cells. Appearance of plaques from AdVHCRGFP/EBV was retarded (by 8 days) compared to non-rearranging viruses, perhaps as a result of basal expression of the liver-specific promoter in 293 cells. However high titer viral stocks of 10^{12} nominal (absorbance-determined) particles/ml was achieved.

Rearrangement in Target and Nontarget Cells

To test excision efficiency, HepG2 (hepatocellular carcinoma) and Hela (cervical carcinoma) cells, obtained from ATCC, were infected with virus at a multiplicity of infection (moi) of 1,000 nominal particles/cell. This titer corresponds to approximately 10 plaque forming units per cell. For these experiments, HepG2 and Hela cells were seeded in 35mm dishes and cultured to approximately 80% confluence in DMEM/FBS as described herein. Cells were infected with the desired multiplicity of virus in a volume of 1 ml at 37 °C for 2 hours. At the end of the incubation, cells were washed with PBS twice and cultured in 2 ml of medium. Cells were collected in parallel at desired points for low molecular weight DNA and RNA extraction. Cells were examined for GFP expression by fluorescence microscopy (Olympus, IX70) or microtiter plate reader (PerSeptive Biosystem, CytoFluor II) before extraction of DNA for analysis of chromosomal rearrangement.

DNA analysis of chromosomal rearrangement was performed as follows. 5 µg of Hirt DNA was digested with Bgl II and analyzed by DNA blot techniques using a labeled EBNA-1 gene fragment as probe (Figure 6). The Bgl II fragment from the non-circularized AdV is 3162 bp, generated from the 5' end of the AdV to the first BglII site in the AdV. The circularized fragment created from the two loxP sites has a size of 4915 bp (Figure 6A). Densitometry revealed that at 72 hours post infection, 95% or more of the input genomes had undergone circularization in HepG2 cells. In contrast, low but detectable levels of circularized fragment was visualized in Hela cells infected at the same time and at the same multiplicity of infection used for the HepG2 cells (Figure 6B).

At the time of infection (t=0, Figure 6B), the amount of input viral DNA detected by DNA blot was higher for HepG2 cells than for Hela cells when similar virus multiplicities were applied (moi of 1,000). This may reflect differences in AdV adsorption or infection efficiency between the two cell types, possibly as a result of the lower levels of coxsackievirus-adenovirus receptor on the Hela cells surface. To achieve similar viral genome input into HepG2 and Hela cells, Hela cells were infected with ten-fold more virus (moi of ~10,000) than HepG2 cells (moi of 1,000). Episomal DNA samples were extracted and analyzed by blotting. The results (Figure 6C) indicated that when comparable amounts of viral genome are present in the nucleus, the cyclization rate in both cell types was similar. Because the level of subsequent GFP expression is

much higher in HepG2 cells than in HeLa cells (Figure 7A), it is likely that very small amounts of Cre recombinase suffice to promote rearrangement, and that recombinase expression is not limiting for rearrangement in either HepG2 or HeLa cells.

5 GFP expression cannot be detected until rearrangement has taken place, so the measurement of the fraction of GFP positive cells provided a simple alternate method for assessing the degree of productive rearrangement. Figure 7A shows that GFP expression developed quickly in transduced HepG2 cells, but that only a few GFP positive cells can be detected in HeLa cells infected with a ten fold higher moi, conditions that allow circularization to a comparable extent to that seen in HepG2 cells
10 (Figure 6C).

The HCR12 promoter specificity was also tested by infecting two additional non-hepatic cell lines, A431 (human epidermoid carcinoma) and HT29 (human colon adenocarcinoma), with the Ad2HCRGFP/EBV vector. Both cell lines were obtained from ATCC and cultured using DMEM/FBS as described herein. A few cells, with
15 weak GFP signal, were detected at 72 hours after infection in these cells (Figure 7B). In contrast, these non-hepatic cells could be infected efficiently with a first generation AdV, Ad2CMVGFP virus (data not shown), indicating that the low GFP signal was not due to the low infectivity of these cells by AdV.

To further assess the utility of the AdV genome rearrangement, primary human
20 hepatocytes were infected with the Ad2HCRGFP/EBV vector. For these experiments, primary human hepatocytes, generously provided by Dr. Albert Edge (Diacrin, Inc., Charlestown, MA) were isolated and cultured as described by Gunsalus et al. (Nat. Med. 3:48-53, 1997), infected with adenovirus, and GFP expression was analyzed. As shown in Figure 7C, GFP expression was readily detected 72 hours after infection.

25

Diminished Viral Gene Expression in Rearranged AdV

After excision, the adenovirus major enhancer/packaging signal segregates with the episomal DNA, yielding a linear fragment containing the remainder of the AdV genome without this important cis-element (Figure 5). To assess the impact of enhancer
30 deletion, PCR amplification and quantitative RT-PCR measurement of late viral gene expression was performed as follows.

Four µg of total RNA was reverse transcribed into cDNA using M-MLV RT by a standard protocol (Promega). 1 µl of the cDNA from each sample was used in subsequent PCR reactions. PCR primers were designed to amplify the tripartite leader sequence of the adenovirus late genes: TPL1 - 5' act ctc ttc cgc atc gct gt 3' (SEQ ID NO.: 7) and TPL2 - 5' ctt gcg act gtg act ggt tag 3' (SEQ ID NO.:8). For detection of the AdV genome in the Hirt DNA samples, 1 µg DNA was employed in the PCR amplification using the following primers which are specific for the adenovirus DNA in the fiber gene: Fiber1 - 5' ccg cac cca cta tct tca ta 3' (SEQ ID NO.: 9) and Fiber2- 5' ggt gtc caa agg ttc gga ga 3' (SEQ ID NO.: 10). PCR reactions were performed as 95 °C 30 seconds; 54 °C 30 seconds; 72 °C 30 seconds for 30 cycles. All amplified products were analyzed on a 2% agarose gel.

For quantitative PCR, a molecular beacon based universal amplification and detection system was used (Intergen). A common leading sequence (Z sequence, 5' act gaa cct gac cgt aca 3') was added to the TPL1 and Fiber1 primers. The TPL2 and Fiber2 primers, described above, were used in the quantitative PCR reactions. 1 µl of the cDNA and one µg of Hirt DNA from each sample were used in the assay. The PCR were carried out in a 96-well spectrofluorometric thermal cycler (Applied Biosystems Prism 7700). The number of template molecules in the PCR reaction was calculated from the standard curve using linearized plasmid as templates.

As most late adenoviral genes transcripts share a common ~200 bp tripartite leader sequence (TPL) (Akusjarvi and Persson, Nature 292:420-6, 1981), the TPL sequence was chosen as a marker of viral gene expression. HepG2 cells were infected with the first generation vectors Ad2CMVGFP and Ad2HCRGFP, or the self-resolving vector, Ad2HCRGFP/EBV, using increasing multiplicities of infection. Total cellular RNA and low molecular weight DNA were isolated in parallel as described by Hirt (J. Mol. Biol. 26:365-9, 1967) and total RNA was prepared using RNazol solution (Tel-Test, Inc.). RT-PCR was performed to quantitate the amount of RNA encoding the TPL in the cDNA samples. PCR amplification of a 201 bp fiber gene fragment from the AdV genome was used to detect the amount of viral genome in the DNA samples. A representative result of three experiments is shown in Figure 8A. TPL sequences were detected, 72 hours post-infection, with either 100 or 1000 viruses infected per cell, using both of the first generation adenoviruses (upper panel).

In contrast, no TPL signal was detected in the self-resolving Ad2HCRGFP/EBV infected cells, even at a moi of 100,000/cell. PCR amplification of the AdV fiber gene revealed comparable levels of AdV genomic DNA in cells infected at comparable moi's. (Figure 8A, lower panel). The cDNA samples in which the TPL signals were detected were further analyzed by real-time fluorescence PCR. The corresponding genomic DNA samples were also analyzed to determine the number of AdV genomes present in each sample. The results are summarized in Figure 8B. There were approximately 1×10^4 TPL per 1×10^6 AdV genomes detected in the Ad2HCRGFP infected cells, but no detectable TPL in the self-resolving Ad2HCRGFP/EBV infected cells. These results indicate that adenoviral gene expression was dramatically reduced by the separation of the viral enhancer sequences occasioned by the re-arrangement of the self-resolving vector.

TYPE A AND TYPE B VECTORS – EXPERIMENTAL RESULTS

Additional experimental examples now follow that further illustrate the general approaches of the invention relating to using and constructing type A and type B vectors. For generating such adenoviral vectors, DNA sequences important for gene expression in the target tissue were placed between two loxP sites. The first loxP site was inserted between the Ad2 left-end inverted terminal repeat (ITR) and the enhancer sequence, replacing a BspLU11I and BstZ17 fragment of Ad2. A target gene expression cassette, comprising a promoter, a gene of interest, polyadenylation signals, the EBV replicon, and site specific recombinase expression unit were inserted in place of the E1 locus.

In type A adenoviral vectors, the second loxP site is placed between TetR and VP16, preserving the coding frame of both (Figure 1A). A bidirectional promoter in which a central heptamer of tetracycline operator sites (TetO) (Gossen and Bujard, Proc. Natl. Acad. Sci. USA 89:5547-5551, 1992) was flanked by two divergently oriented basal elements, directs the expression of TetR loxP VP16 from a synthetic TATA element, whereas Cre recombinase is controlled by the same heptamer of operator upstream of the HIV LTR basal element.

In the case of type B viruses (Figure 1B), the second loxP site was inserted in the first intron of the Efl α gene, which contains the transcription stimulating sequences described herein. In addition, a splice acceptor sequence was added to the 5' end of the coding sequence of the gene of interest. To avoid rearrangement during plasmid

construction in bacteria, the Cre recombinase coding sequence was interrupted by the addition of the human IgG1 hinge-CH2 intron (between amino acids Q78 and A79), as described herein.

5 Designing a Compact EBV Replicon

Most plasmids employing the EBV latent origin of replication exceed 10 kb in length. To provide a means for increasing the capacity of the recombinant adenoviral type A or type B vectors to accommodate a therapeutic gene, a compact EBV replicon having episomal stability was designed. To this end, deletions were generated in both
10 the cis-acting origin of replication, OriP, and the sequences encoding the trans-acting replication protein, Epstein Barr virus nuclear antigen-1 (EBNA-1) (Figure 9A). Episomal persistence was assessed with a green fluorescent protein (GFP)-bearing test plasmid by determining the fraction of cells retaining green fluorescence as a function of time, assuming that the half-life of GFP, in daughter cells that have not received an
15 episome as a result of segregation failure, is approximately 1.4 days (Fukumura et al., Cell 94:715-725, 1998).

EBNA-1 contains a central repeated structure that consists entirely of Gly and Ala residues, termed the GA repeats (Figure 9A). Although deletion of this structure has been reported to have little consequence, a deletion mutant consisting of both a short
20 OriP and a short EBNA-1 (SoriP + SEBNA1) was generated and found not to support plasmid maintenance effectively (approximately 40% loss per cell division). A version of this mutant, reconstructed with 40 GA repeats, in which the short OriP was paired with a short EBNA-1 provided significantly better plasmid stability (20% loss per cell division vs. 10 % per cell division for the wild type) (Figure 9B). Since most target
25 tissues are relatively quiescent mitotically, this level of segregation fidelity provides reasonable stability in a compact replicon.

Producing Cell Lines that Express Cre- or FLP-Dominant Negative Mutants

As discussed herein, one obstacle to creating adenovirus carrying both
30 recombinase and target sites has been the difficulty of controlling recombinase activity during virus propagation. Since efficient recombinase activity is needed in target cells, recombinase activity is best tempered in the production cell line.

Vector-independent methods to suppress recombinase activity during the production phase are attractive because they allow vector design objectives to be pursued with fewer constraints. In principle, dominant negative recombinase mutants provide the desired antagonism of recombinase activity. Cell lines expressing such
5 recombinase dominant negative mutants were produced as follows.

Dominant negative Cre mutants were selected from known point mutants (Wierzbicki et al., J. Mol. Biol. 195:785-794, 1987) that are defective in recombination function but are likely to retain dimerization function Figures 10A-E. Several mutants were screened for their abilities to inhibit Cre activity of a type B vector construct
10 (ad2239 in Figure 10B) in a transient cotransfection assay. Under these conditions, Cre activity is detected by the expression of GFP that occurs upon rearrangement. Figures 10D and 10E show the point mutants that were assessed and their relative activities in the transient cotransfection assay. Dilution studies, in which increasing amounts of substrate/Cre plasmid were cotransfected with the mutant forms, were conducted and
15 based on its favorable profile, one mutant recombinase, designated CreY324C, was chosen for further development (Figure 10D).

Strong constitutive expression of CreY324C, under control of the Efl α promoter failed to yield stable cell lines. Stable clones were obtained when the Efl α promoter was replaced with a tetracycline regulated promoter (Gossen et al., Science 268:1766-
20 1769, 1995). Clones were then tested for the ability to inhibit Cre enzymatic activities, and one clone, designated cell line #17, was selected for additional experiments. When a plasmid bearing Cre and capable of undergoing Cre-directed rearrangement to create a GFP transcription unit (ad2239) was transfected to #17 cells or parental 293ON cells, GFP expression in the #17 cells in the presence of 2 μ M doxycycline was significantly
25 lower than those of controls (Figure 11), showing that Cre enzyme activity can be inhibited in #17 cells.

In addition to dominant negative Cre mutants, dominant negative FLP mutants may also be identified. FLP belongs to the same family of site-specific recombinases as Cre recombinase. A number of FLP mutations that show defects in either cleavage or
30 ligation of FRT sites have been identified. Mutant FLP defective in cleaving FRT site (for example, H309L, L315P, G328R, G28E, N329D, S336Y, S336F, A339D, Y343F, and H345L) are generated using standard methods. Mutants that inhibit the wild type enzyme are then identified for generating stable cell lines according to the methods

described above. These and the other cell lines (described herein) are then used for producing FRT/FLP containing virus.

As mentioned above, difficulties creating stable cell lines expressing Cre dominant negative mutants were occasionally encountered. This difficulty is not limited to Cre mutants, but also to the wild-type Cre enzyme. In contrast, 293 cell lines stably expressing a thermostable FLP, referred to as FLPe (Buckholz et al., Nat. Biotechnol. 16:657-662, 1998), were created, suggesting that FLPe might not be as cytostatic as Cre protein. To demonstrate this, 293 cells were transfected with plasmids expressing either Cre or FLPe, and puromycin resistant colonies were selected. To generate stable cell lines expressing Cre or FLP mutants, 293 TetON cells were transfected with linearized plasmid expressing Cre or FLP mutants and puromycin acetyltransferase and selected with 1 µg/ml of puromycin. Puromycin resistant colonies were characterized further for their ability to inhibit Cre recombinase using the cre (ad2239) or flp (ad2879) substrate plasmids. Table 1 shows that there are more puromycin resistant colonies selected from FLPe transfected cells than from Cre transfected cells. From this result, it is expected that stable cell lines expressing a reasonably high level of dominant negative FLP may be readily created.

TABLE 1

Puromycin Resistant Colonies Formed When Cre Expressing or FLPe Expressing Plasmid was Used to Transfect 293 Cells

Expression Plasmid	Number of colonies (2 µg/ml puromycin)
Control	236
Cre	92
FLPe	127

Cre or Cre dominant negative mutants were also found to inhibit FLP activity (Table 2). Accordingly, cell lines such as cell line #17, that stably express a Cre dominant negative mutant (for example, CreY324C), are useful for producing FLP/FRT carrying adenovirus.

TABLE 2

Cre Inhibition of FLP Activity *in trans*

Plasmids	Arbitrary GFP intensity
Efl α FLP + FLP substrate + vector control	4.9
Efl α FLP + FLP substrate + Efl α Cre	0.78
Efl α FLP + FLP substrate + Efl α Cre R173C	2.23

FLP enzyme activity was measured by the GFP intensity by cotransfecting with a FLP substrate plasmid, ad2879 (Figure 19A). GFP intensity was quantified using IP lab software.

10 Transcriptional Regulation of Cre or FLP Recombinases

It has been relatively difficult to achieve high-level promoter inducibility in a replicating adenovirus. The challenge is similar to that of achieving faithful control of transcription in a transient expression setting. One approach to increase the induction ratio in a transient setting is the use of auto-regulatory (feed-forward) circuits. One such system, based on tetracycline dependent activation, is shown in Figure 12. A central heptamer of tetracycline promoter operator sites (TetO sites) was placed between two divergently oriented basal TATA elements. The leftward TATA controls the expression of the TetR-VP16 fusion protein, in which a loxP (or FRT) site has been placed between the TetR DNA binding domain and the VP16 transcriptional activator. The rightward TATA box directs the synthesis of recombinase, either Cre or the yeast FLP enzyme. In the presence of tetracycline, the promoter has reduced activity in both directions. Upon removal of tetracycline, the synthesis of both TetR-VP16 and recombinase are induced (Figure 1A). The induced recombinase then disjoins the TetR DNA binding element from the transcriptional activation contributed by VP16. Any existing TetR-VP16 fusions thereafter promote transcription of TetR, which competes with TetR-VP16 for TetO, resulting in deinduction of recombinase transcription.

When a model target cell line, HepG2, was tested with this type of adenovirus, the efficiency of circularization was low relative to that seen in 293 cells (data not shown), indicating a cell dependence of the bidirectional TetO promoter. To correct this, the TATA element of the TetO synthetic promoter (derived from the CMV immediate early promoter) was replaced with that of the HIV LTR. Constructs bearing differing components of the HIV basal promoter were analyzed for strength and regulation in 293 and HepG2 cells (Figure 13A). Among the constructs tested, one

version bearing the HIV LTR TATA and Sp1 elements (D in Figure 13A) showed the least basal expression in 293 cells (data not shown) and the greatest induction in HepG2 cells (Figure 13B).

Using this promoter, a construct (ad3400) containing the autoregulatory structure D as shown in Figure 13A was engineered, and Cre activities in the presence and in the absence of tetracycline were assayed. Plasmid ad2265 (Figure 14) in which a blue fluorescent protein (BFP) expression unit is interrupted by two loxP sites and transcription termination sequences was used as a substrate for Cre. Cre-mediated recombination joins BFP to the promoter resulting in BFP expression. As shown in Table 3, no difference was found in the intensity of BFP expression, either in the presence or absence of tetracycline. One possible explanation for this is that very little Cre protein is required for activity. Consistent with this idea, standard immunohistochemical techniques failed to reveal the presence of Cre enzyme in cells that were fully induced (data not shown).

TABLE 3

Cre Recombinase Activity Regulation in Type A Constructs

Construct	Cre Form	+tet -tam	-tet -tam	+tet +tam	-tet +tam
ad3400	Cre	2.58	3.39	2.71	6.20
ad4394	Cre-LBD	0.081	0.22	0.056	1.02
ad4705	LBD-Cre-LBD	ND	ND	ND	ND

Cre enzyme activity was measured in the presence or absence of the ligands, tamoxifen, by cotransfecting with the substrate plasmid (ad2265). BFP intensity (mean intensity/area) was quantified by analyzing fluorescent images captured by a digital camera using IP lab software. ND, refers to fluorescent intensities that were too weak to measure. Tet, Tetracycline; tam, tamoxifen.

Deletion of the PolyA Consensus Sequence from Cre or FLP Transcription Units

To reduce the expression of FLP or Cre recombinase further, the consensus polyA addition signals from the Cre or FLP transcript unit were deleted from vector constructs, leaving polyadenylation dependent on distal downstream sequences, for example, in gene IX. The activity of Cre using type B proviral constructs with or without the polyA signal was measured. As shown Table 4, the construct without polyA signals (AD229.3) showed a significant reduction of GFP intensity compared to a

construct bearing the polyA signal (AD230.5). When FLPe constructs of similar structure were evaluated, similar results were found (data not shown). These data show that Cre and FLPe enzyme activity levels can be modulated by attenuating polyadenylation.

5

TABLE 4

Effect of Deleting polyA Addition Signal From the Cre Expression Unit on Cre Enzyme Activity Level

10

	polyA	Relative GFP Intensity
AD229.3	-	0.25
AD230.5	+	1

Post-Transcriptional Regulation of Cre Recombinase Activity

Post-transcriptional control mechanisms of Cre recombinase activity were also evaluated. Translational fusions between Cre and the ligand binding domain (LBD) of estrogen receptor have been reported to be regulated by estrogen (Feil et al., Proc. Natl. Acad. Sci., U.S.A. 93:10887-10890, 1996; Gossen et al., Proc. Natl. Acad. Sci., U.S.A. 89:5547-5551, 1994), or, in the case of mutant estrogen receptors (Metzger et al., Proc. Natl. Acad. Sci. U.S.A. 92:6991-6995, 1995), by the partial antagonist tamoxifen.

Use of a ligand-dependent recombinase (ad4394 in Table 3), in combination with the HIV LTR-based autoregulated Tet system, allowed for a small degree of regulation by tetracycline, but not by ligand, as assayed using the ad2265 rearrangement assay (Table 3). One interpretation of this finding is that fusion of the estrogen receptor LBD to Cre provides only modest control of recombinase activity, but attenuates enzyme potency to a level so that transcriptional regulation can be measured.

To increase control of recombinase activity, the LBD was fused both to the N-terminus and C-terminus of Cre (LBD-Cre-LBD) and inserted into the coding sequence of both type A and type B vectors. When the LBD-Cre-LBD construct of type A was transfected into 293 cells, it showed no significant Cre enzyme activity even in the presence of ligand (Table 3). This result confirmed that the Cre recombinase activity is attenuated by N-terminal or C-terminal extension.

When the LBD fusion Cre enzymes were assayed in the type B vector context, only LBD-Cre-LBD fusions (pk8-ad4626) showed ligand-dependent regulation of Cre enzyme activities (Table 5). It appears that attenuated Cre activity in LBD-Cre-LBD, in the absence of ligand, is low enough to fall below the upper limit of the Cre assay.

5

TABLE 5

Cre Enzyme Activities of Type B Provirus

Cre Form	Provirus	-tam	+tam
Cre	pk8-ad2239	ND	ND
Cre-LBD	pk8-ad4332	4.1	6
LBD-Cre-LBD	pk8-ad4626	0.05	4.2

10

Cre enzyme activity was measured in the presence or absence of the ligands, tamoxifen. GFP intensity was quantified using IP lab software.
tam, tamoxifen.

15

Consistent with this notion, only the construct carrying two LBDs, pk8- ad4626, was able to produce virus (AD121.5) by transfection and propagate in 293 cells, while pk8-ad4332, which carried one LBD, produced virus (AD100.9) initially (following transfection of the cognate DNA) but was unable to propagate in 293 cells (Table 6). In the case of wild type Cre, no virus was produced in 293 cells by transfection.

20

TABLE 6

Production and Propagation of Type B Adenovirus

25

Cre Form	Type B adenovirus	Viral Production in 293 cells	Viral Propagation in 293 cells	Viral Propagation in #17 cells
Cre	Pack8-2239	-	-	-
Cre-LBD	AD100.9 (Pack8-4332)	+	-	+
LBD-Cre-LBD	AD121.5 (Pack8-4626)	+	+	+

30

The AD100.9 virus was able to propagate in #17 cells expressing the dominant negative Cre Y324C, demonstrating that modulation of Cre activity is important for viral production. Thus, adenovirus carrying both two loxP sites and Cre in two different configurations were generated by controlling Cre activity.

5

Viral Rearrangement in Culture

Cre/loxP mediated rearrangement of the adenovirus in tissue culture cells has also been analyzed. As shown in Figure 15A, the AD121.5 virus showed a significant increase in GFP expression in the presence of the ligand, estrogen, suggesting a successful rearrangement of the virus by Cre recombinases. When non-chromosomal DNA (Hirt, *J. Mol. Biol.* 40:141-144, 1969) was made from the cells and analyzed by DNA blot analysis, the viral DNA from estrogen treated cells was identified mostly in circular form (C in Figure 15B), while the DNA from cells not treated with estrogen was found mainly in linear form (L in Figure 15B).

To evaluate the efficiency of the self-rearranging viruses *in vivo*, high titer stocks of AD102.7 (a type A virus carrying LBD-Cre, pk8-ad4394) in #17 cells was prepared and purified by CsCl gradient ultracentrifugation. The titer of AD102.7 ($4-6 \times 10^{12}$ /ml by OD) is comparable to or slightly exceeds that of control viruses ($2-4 \times 10^{12}$ /ml by OD) which carry neither Cre nor a loxP site. To determine the efficiency of viral rearrangement *in vivo* and whether such rearrangement is dependent on the presence of ligand, AD102.7 virus (4×10^{11} pfu/mouse as determined by optical density) were injected via tail vein into Rag-2 mice that were pretreated with vehicle alone or 110 μ g/day of tamoxifen for 7 days as follows.

Rag-2 mice were injected with either PBS (mock) or 4×10^{11} adenovirus particles (as determined by OD₂₆₀) of type A virus, AD102.7, via the tail vein. At various times after injection, animals were sacrificed and the liver tissues were removed and frozen rapidly on dry ice. To visualize GFP expression in animal tissues, mice were anaesthetized and perfused with 4% paraformaldehyde containing 0.2% glutaraldehyde intracardially (Kafri et al., *Natl. Genet.* 17:314-317, 1997), and the liver tissues were removed and fixed overnight at room temperature in the perfusion buffer containing 30% sucrose. The fixed tissues were sectioned serially and observed under confocal scanning laser microscopy. In experiments evaluating the responses of ligand-regulated recombinase, mice were injected either with vehicle (vegetable oil) alone or with 110 μ g/day of tamoxifen for 7 days prior to adenoviral injection.

Liver tissues from these animals were harvested at 2.5 hrs post injection (the earliest time point taken after injection) and Hirt DNA from approximately 250 mg of frozen hepatic tissue was prepared and analyzed by blot analysis. As shown in Figure 16, the majority of adenoviral DNA was found in circular form in tissues from untreated mice, as well as tamoxifen-treated mice. It can be concluded from these data that the Cre enzyme activity present in the tissue, even in the absence of ligand was sufficient for efficient self rearrangement of virus. As expected, the hepatic tissues from the Rag2 mice injected with AD102.7 showed strong expression of GFP (Figure 17).

Demonstrating that AD102.7 virus, produced efficiently in 293 cells at high titres by the conventional means, can self rearrange efficiently *in vivo* provides the proof of the concept that potentially safer adenoviral gene therapy vectors can be produced.

Adenoviruses Carrying Both FRT and FLP Recombinase

Type A and type B proviral constructs carrying both FRT (FLP recombinase recognition site) and FLP recombinase were also generated. Structures of these viruses are analogous to those of loxP/Cre carrying viruses except that loxP sites are replaced by FRT sites and Cre coding sequence is replaced by FLP coding sequence (Figures 18A, 18B).

Virus Production at Reduced Temperature

Temperature dependence of the Efl α promoter using GFP expression as a marker was also examined. As shown in Table 7, Efl α promoter activity is strongly reduced at 32°C in comparison to 37°C or 39°C. The temperature sensitive nature of the Efl α promoter was used to propagate type B adenovirus carrying FLP at 32°C following initial production of the virus by DNA transfection (pk8-ad3302) at 37°C. HepG2 cells infected with these viruses (AD41.4) showed strong GFP expression, but with an approximately 12 hr delay compared to GFP expressing viruses, suggesting that FLP recombinase activity may be impaired at 37°C. To improve the activity of FLP recombinase, viral constructs were created using a thermostable FLP (referred to as “FLPe”) described by Buchholz *et al.* (Nat. Biotechnol. 16:657-662, 1998).

TABLE 7

Effects of Temperature on Efl α Promoter Strength as Shown by GFP Intensities

Tester plasmids	Arbitrary GFP Intensities			
		32°C	37°C	39°C
Efl α GFP	16 hrs	1475	7886	11409
	41 hrs	6472	36699	50787
	86 hrs	16256	53370	54424
Efl α Cre +ad2204	16 hrs	243	1141	2132
	41 hrs	1094	9119	9784
	86 hrs	695	3219	8144

5 GFP intensities were measured using a Fluorescent reader.

The activities of FLP and FLPe using a FLP substrate plasmid (ad2879, Figure 19A) in 293 cells were compared. As shown in Table 8, FLPe is significantly more
 10 active than FLP under these conditions.

TABLE 8

FLPe is Significantly More Active than FLP Recombinase

15

Plasmid		Mean GFP intensity
ad4821	Efl α FLPe	2.39
ad2949	Efl α FLP	0.01

Plasmid coding either FLPe or FLP was cotransfected with a FLP substrate plasmid (Figure 19A) into 293 cells. GFP intensity of each transfection was measured using IP lab program.

20

In addition, a tamoxifen-regulated FLPe was created by fusing the ligand-binding domain from a mutant form of estrogen receptor to the FLPe coding sequence at its C-terminus (FLPe-LBD). The FLPe-LBD was found to be regulated by the ligand, tamoxifen (Table 9). Although FLP activity was retained by C-terminal fusion (FLP-
 25 LBD), addition of a short oligopeptide tag to the N-terminus of FLP abolished its activity (data not shown).

TABLE 9

Tamoxifen Regulation of FLPe as Determined by GFP Intensities

Plasmid	FLPe	-tam	+ tam
ad4821 +ad2879	FLPe	++++	++++
ad5022 +ad2879	FLPe-LBD(tam)	+	+++

GFP intensity resulting from FLPe mediated recombination was measured using a fluorescent microscope. tam, 2 µg/ml tamoxifen.

10 Inhibition of FLPe Activities by Anti-sense FLPe

An anti-sense approach to inhibit FLP enzyme activity was also employed. This approach tested the notion that incorporation of an open reading frame into an antisense transcript would stabilize the transcript and potentiate antisense activity. Two approaches were utilized. In one approach, the BFP coding sequence was placed upstream of anti-FLPe. In the second approach, an anti-FLPe was placed upstream of an internal ribosome entry sequence (IRES) and the BFP coding sequence (Figure 20). The ability of these constructs to inhibit FLPe was assayed using a FLP substrate plasmid, ad2879 (Figure 19A) and the result is shown in Figure 21. These data show that anti-sense FLPe is more effective in inhibiting FLPe function when it is fused to BFP, which can presumably be replaced with any other stable protein.

OTHER SELF-REARRANGING ADENOVIRUSESMixed Infection With Adenoviruses Carrying loxP/FLP and FRT/Cre

One of the ways to produce adenoviruses that can be rearranged in target cells but not producer cells is to engineer two separate viruses, each carrying one recombinase and the target sequence for the other. To test this system, type B adenoviral constructs carrying Cre recombinase and FRT sites and FLPe recombinase and loxP sites were created (Figure 22). In target cells infected with both viruses, Cre catalyzes recombination between the two loxP sites in the FLP virus, and FLP carries out FRT mediated recombination in the Cre virus, resulting in two circular plasmids. The loxP virus contained BFP, whereas the FRT virus contained GFP. Measurement of the fluorescent intensities of GFP and BFP, after cotransfecting the two constructs, revealed that BFP expression (mediated by Cre enzyme) was greater than GFP expression

(mediated by FLP enzyme), suggesting that the Cre enzyme functions more efficiently than FLP.

Accordingly, these two recombinase activities in the target cells need to be balanced for complete circularization of both viral vectors. Exemplary methods for modulating Cre /FLP activity include the use of transcriptional regulation (such as by varying promoter strength and/or with or without poly A addition signal sequence) and translational and/or post-translational regulation (such as by changing FLP to FLPe and making LBD fusion proteins), and post viral production control (such as by changing the ratio of two viruses).

In one approach, Cre was replaced by Cre-LBD and FLP was replaced by FLPe. To improve identification of the rearrangement products, BFP was replaced with RFP as a marker for Cre recombination. As shown in Table 10, in the presence of estrogen, expression of GFP (FLPe mediated) and RFP (Cre mediated) were similar.

TABLE 10

RFP and GFP Expression of Cells Cotransfected With Type B Proviral Constructs Carrying Cre-LBD/FRT(GFP) or FLPe/loxP (RFP)

Plasmids	Genotypes	GFP intensity		RFP intensity	
pk8-ads120 +pk8-ads113	Cre-LBD/FRT (GFP) + FLPe/loxP (RFP)	Estrogen			
		-	+	-	+
		3222	3183	46	1954

To insure that both Cre and FLP carrying viruses with an optimal ratio infect each target cell, these viruses can be cross-linked prior to infection. For example, Cre carrying virus is labeled by biotin while FLP carrying virus is labeled by avidin. Mixing two types of modified viruses generates virus complexes of desired proportions as well. Biotinylation or avidinylation can be carried out using commercially available reagents such as EZ-Link TFP-PEO biotin (Pierce) and EZ-Link maleimide activated NeutrAvidin (Pierce). The extent of the biotin/virus and avidin/virus will be empirically determined to ensure the viability of the virus and to obtain an optimal ratio of two viruses in the complex. Optimal ratios will be those resulting in 1:1 Cre and FLP recombinase activities in target cells. The modifications will be done following manufacture's instructions.

This approach not only increases the effective capacity of adenoviral vector but also opens new avenue of applications involving multiple proteins, some of which cannot be coexpressed in production cell line as a result of combination toxicity.

- 5 All references mentioned herein are hereby incorporated by reference.
Other embodiments are within the claims.

What is claimed is:

10

Claims

- 5 1. A replicatable viral DNA vector encoding a site-specific DNA-altering enzyme and a DNA target recognized by said enzyme, said enzyme selectively converting, in a cell expressing said enzyme, said DNA vector to a rearranged form.
- 10 2. The vector of claim 1, wherein said rearranged form comprises an autonomously replicating episome.
3. The vector of claim 1, wherein said rearranged form comprises linear and circular DNAs.
- 15 4. The vector of claim 1, wherein said vector comprises adenoviral DNA.
5. The vector of claim 1, wherein said vector comprises a genetically-engineered recombination site.
- 20 6. The vector of claim 5, wherein said recombination site comprises a target of Cre or FLP.
7. The vector of claim 1, wherein said enzyme comprises a recombinase or an integrase.
- 25 8. The vector of claim 7, wherein said recombinase is Cre or FLP recombinase.
9. The vector of claim 1, wherein said enzyme is functional in a mammalian cell.
- 30 10. The vector of claim 5, wherein said recombination site comprises a recognition sequence of a site-specific DNA-altering enzyme.

11. The vector of claim 1, wherein said vector comprises an origin of replication functioning in a mammalian cell.

5 12. The vector of claim 11, wherein said origin of replication is an Epstein Barr Virus replicon.

13. The vector of claim 1, wherein said vector comprises a gene of interest.

10 14. A method for assembling a recombinant adenoviral DNA said method comprising the steps of: (a) providing a first linearized DNA vector comprising a restriction site and a cos site and a second linearized DNA vector comprising said restriction site, an adenoviral nucleic acid molecule, and a cos site; and (b) ligating said first and second linearized DNA vectors, said ligation assembling a recombinant adenoviral DNA.

15 15. The method of 14, wherein said first linearized DNA vector comprises a selectable marker.

20 16. The method of claim 14, wherein said first linearized DNA vector comprises an adenoviral left end-inverted terminal repeat.

17. The method of claim 14, wherein said first linearized DNA vector comprises a gene of interest.

25 18. The method of claim 14, wherein said second linearized DNA vector comprises a selectable marker.

19. The method of claim 14, wherein said second linearized DNA vector comprises an adenoviral right-end inverted terminal repeat.

30 20. The method of claim 14, said method further comprising packaging said assembled adenoviral DNA into a phage and infecting a host cell.

21. The method of claim 14, wherein said first and second linearized DNAs comprise a cosmid vector.

5 22. The method of claim 14, wherein said adenoviral DNA is flanked by cleavage sites.

23. The method of claim 22, wherein said cleavage sites comprise intron endonuclease cleavage sites.

10 24. An adenovirus producer cell comprising a nucleic acid molecule that expresses a dominant negative site-specific DNA-altering enzyme.

25. The producer cell of claim 24, wherein said site-specific DNA altering enzyme is a *dominant negative recombinase*.

15 26. The producer cell of claim 25, wherein said recombinase is a Cre or Flp recombinase.

20 27. The producer cell of claim 26, wherein said dominant negative recombinase is CreY324C.

28. The producer cell of claim 26, wherein said Flp recombinase is Flpe.

25 29. The producer cell of claim 24, wherein said cell is a 293 human embryonic kidney cell.

30 30. A vector comprising, in the 5' to 3' direction,
a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme;

a gene of interest;

a lineage-specific gene promoter;

a second genetically engineered *cis*- acting target recognized by a site-specific DNA altering enzyme; and

a nucleic acid molecule encoding a site-specific DNA altering enzyme.

35

31. A vector comprising, in the 5' to 3' direction,
a first genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme;
a gene of interest;
5 a bi-directional promoter, comprising a second genetically engineered *cis*-acting target recognized by a site-specific DNA altering enzyme; and
a nucleic acid molecule encoding a site-specific DNA altering enzyme.

32. A method of gene therapy comprising the administration to a patient in
10 need of gene therapy a therapeutically effective amount of the vector of any one of claims 1, 30, or 31 which is expressed in said patient

33. A population of cells transfected with the vector of any one of claims 1,
15 30, or 31.

34. A method of gene therapy comprising the administration to a patient in
need of gene therapy a therapeutically effective amount of the population of cells
of claim 33.

20

1/40

FIG. 1A

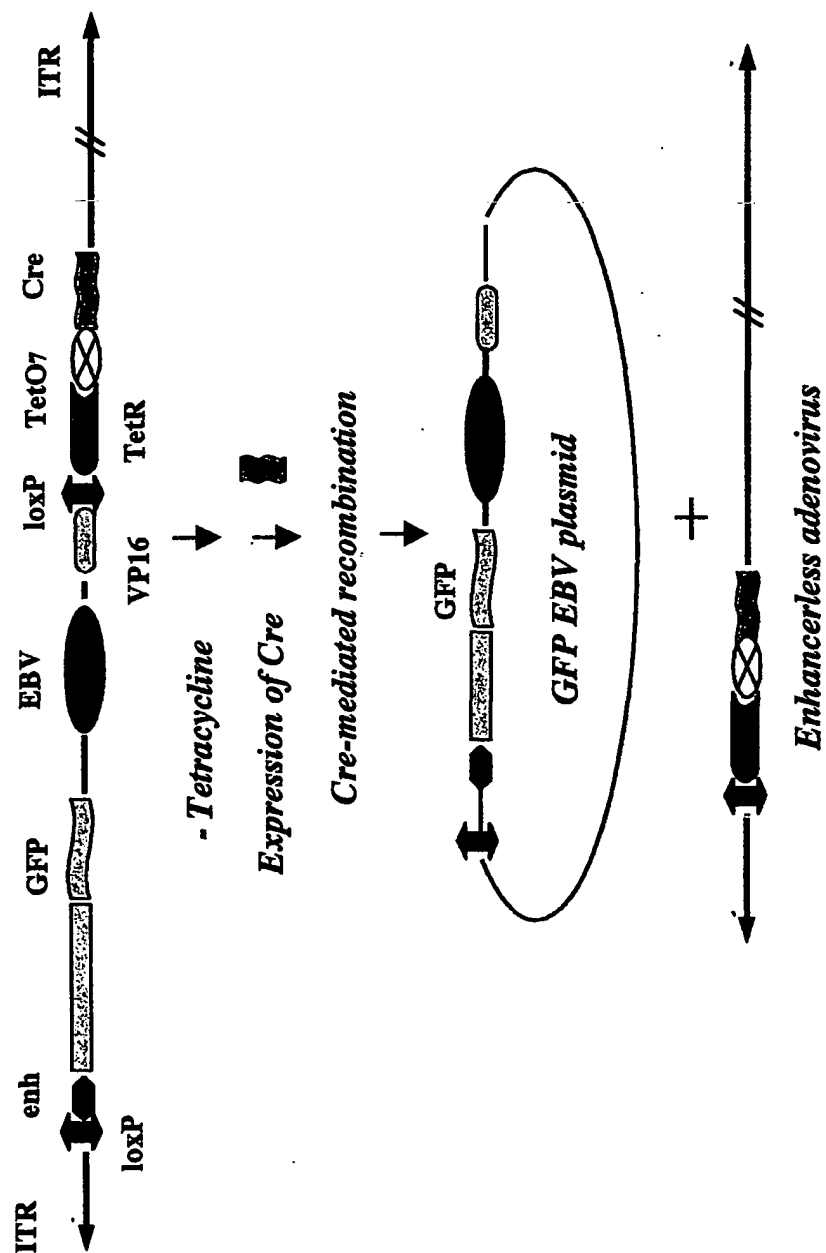
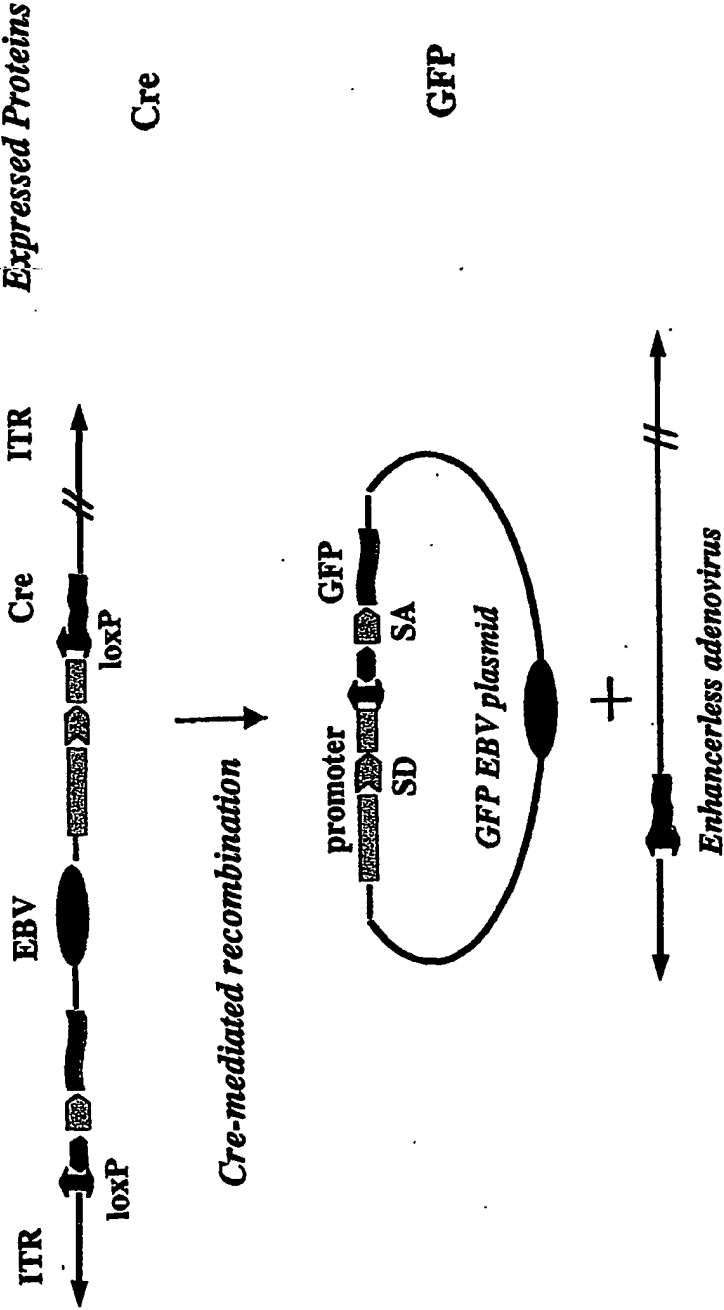
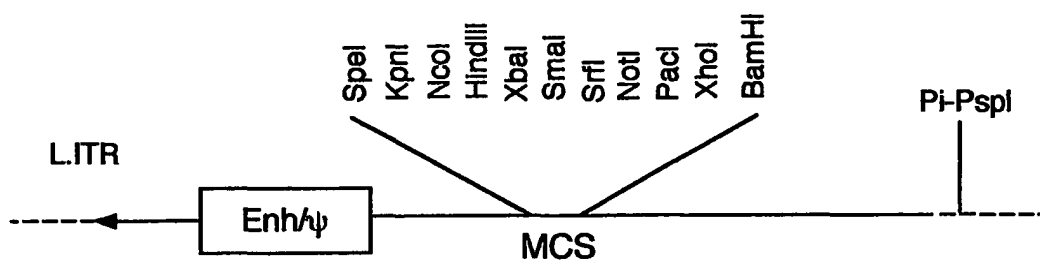


FIG. 1B



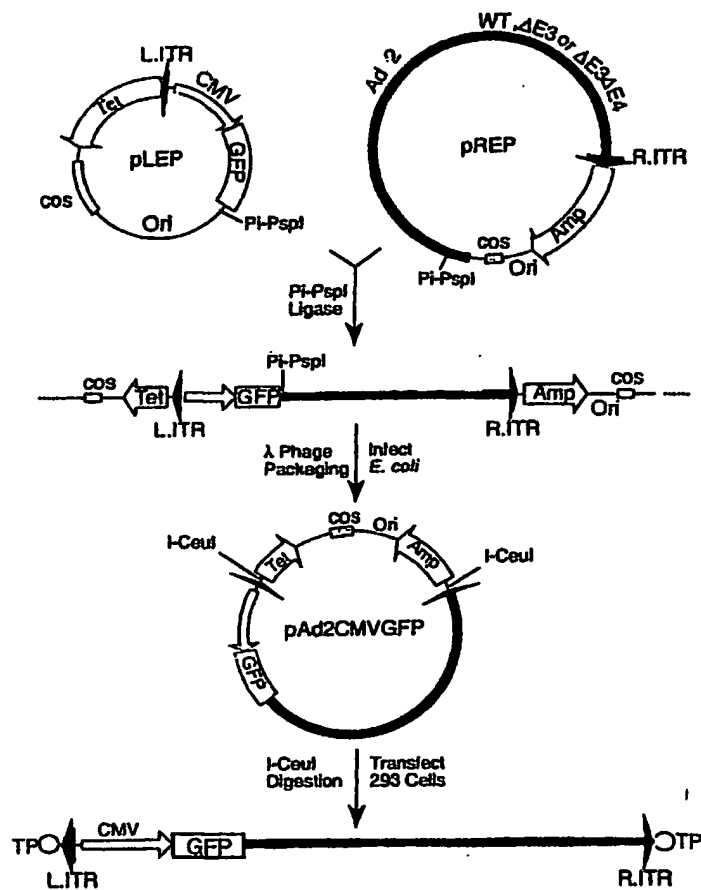
3/40

FIG. 2A



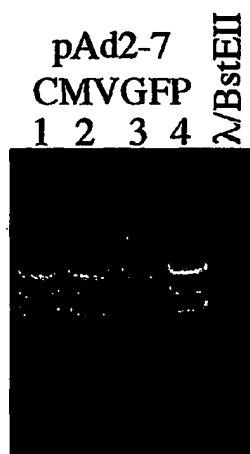
4/40

FIG. 2B



5/40

FIG. 3A



6/40

FIG. 3B

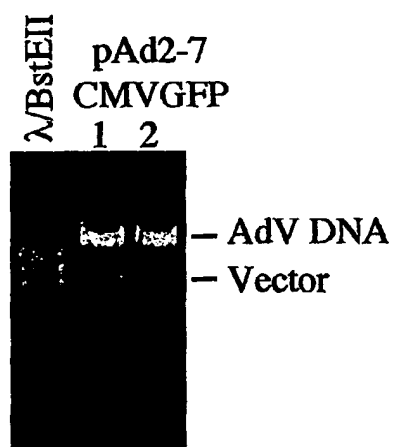
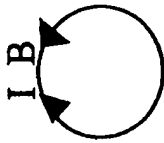
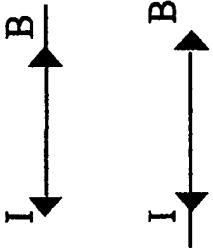


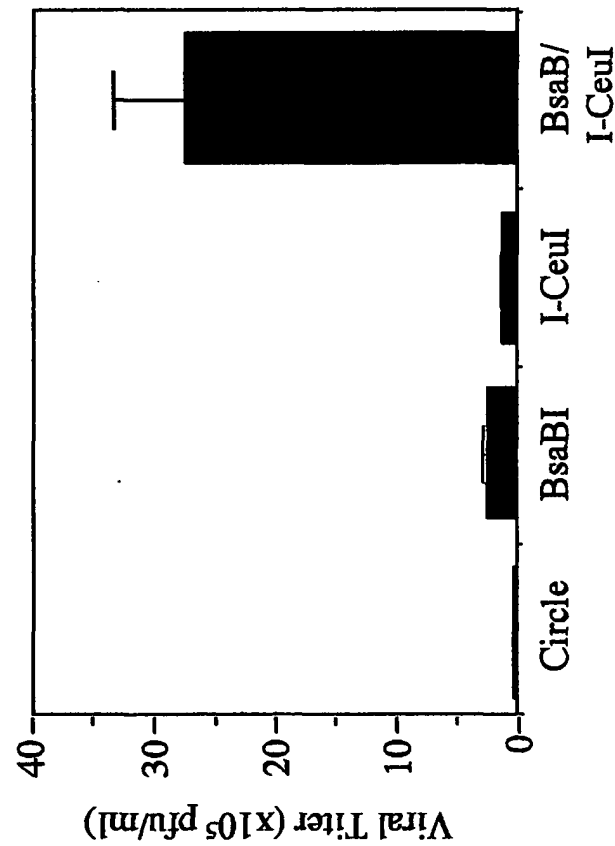


FIG. 4A

Ad DNA forms	Digestion	CPEs	Time (days)
1. Circular Ad DNA		pIAAd2B 0.17	--
2. Linearized cosmid DNA		pIAAd2B / I-CeuI 8.5	10
		pIAAd2B / BsaBI 8.5	10
3. Liberated AdV DNA from cosmid		pIAAd2B / I-CeuI + BsaBI 76	6

8/40

FIG. 4B



9/40

FIG. 5

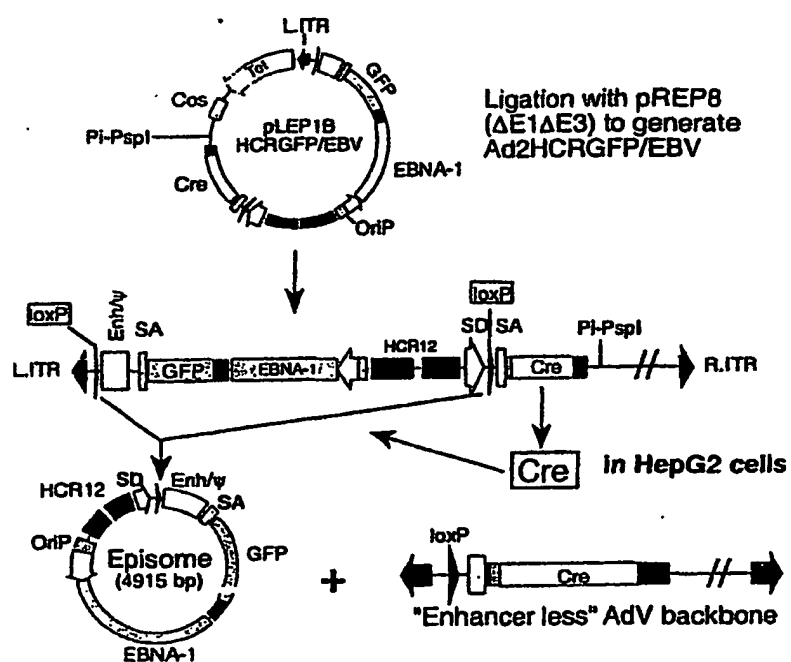
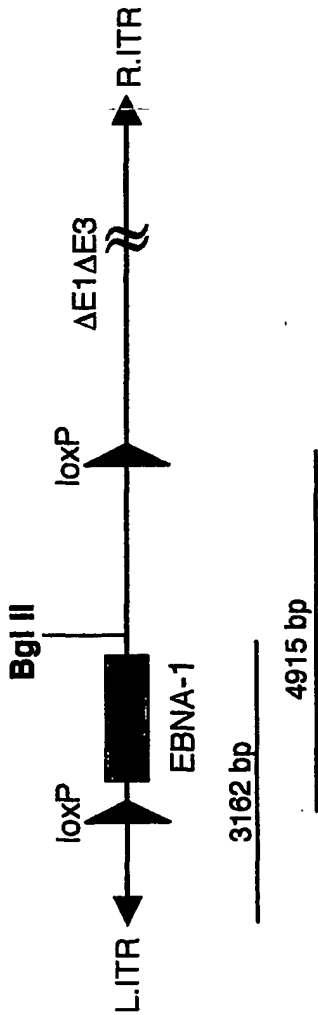
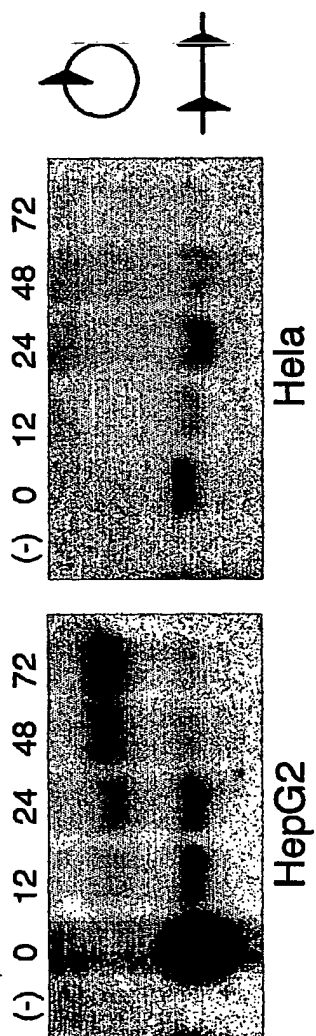


FIG. 6A



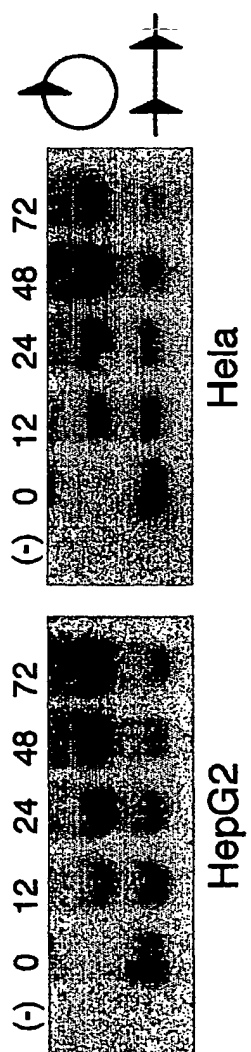
11/40

FIG. 6B



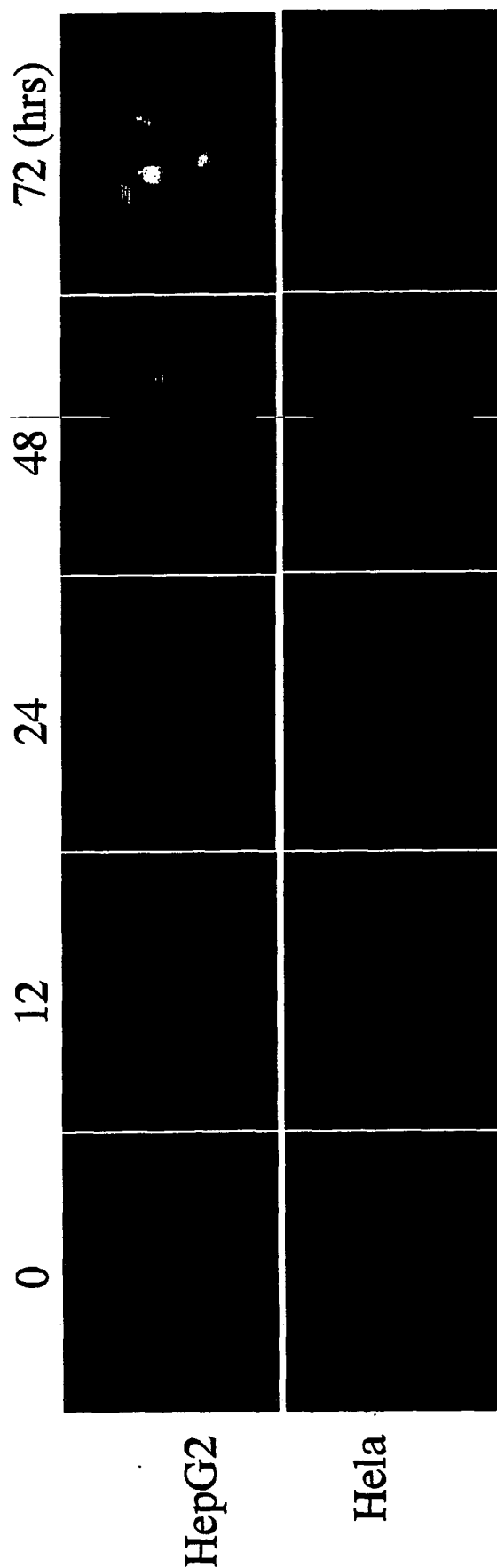
12/40

FIG. 6C



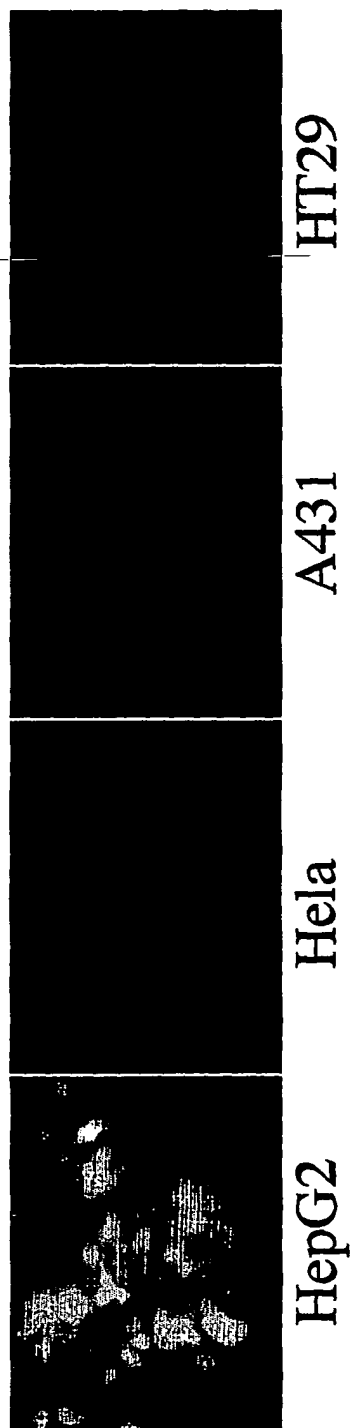
13/40

FIG. 7A



14/40

FIG. 7B



15/40

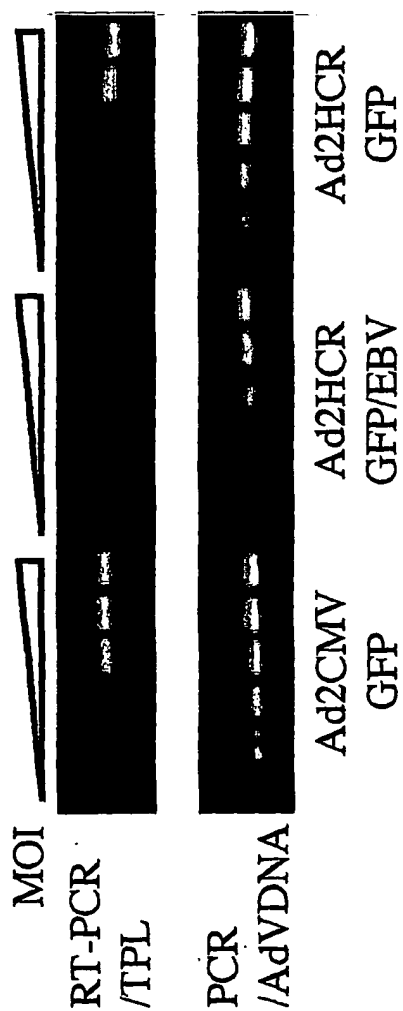
FIG. 7C



Human primary hepatocytes

16/40

FIG. 8A



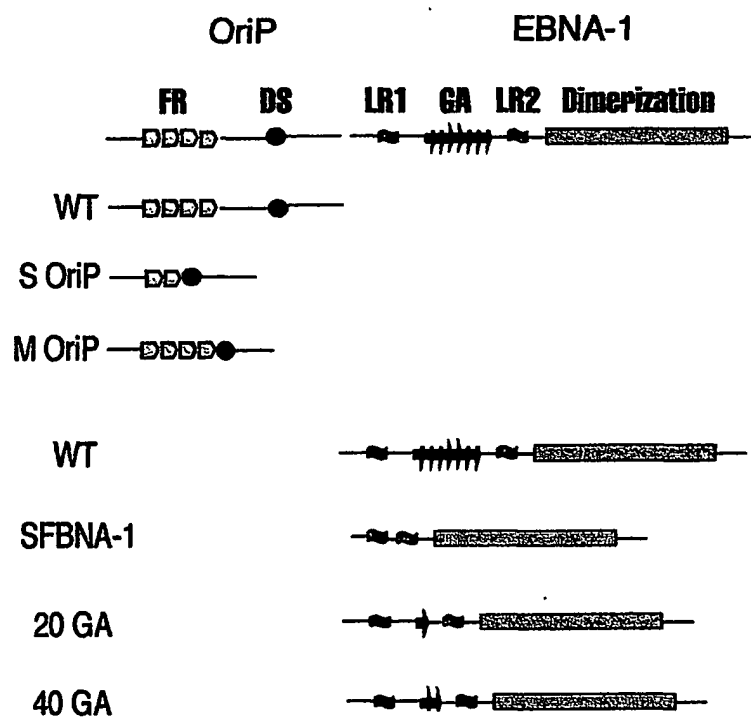
17/40

FIG. 8B

# of TPL mRNA / 10^6 Adv DNA	
Ad2HCRGFP	$1.15 \pm 0.28 \times 10^4$
Ad2HCRGFP/EBV	Non-detectable

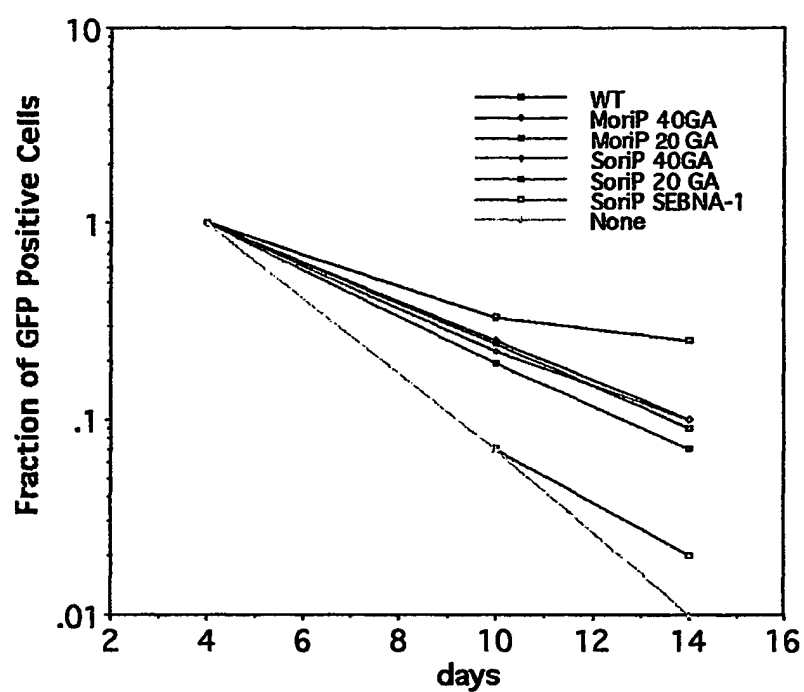
18/40

FIG. 9A



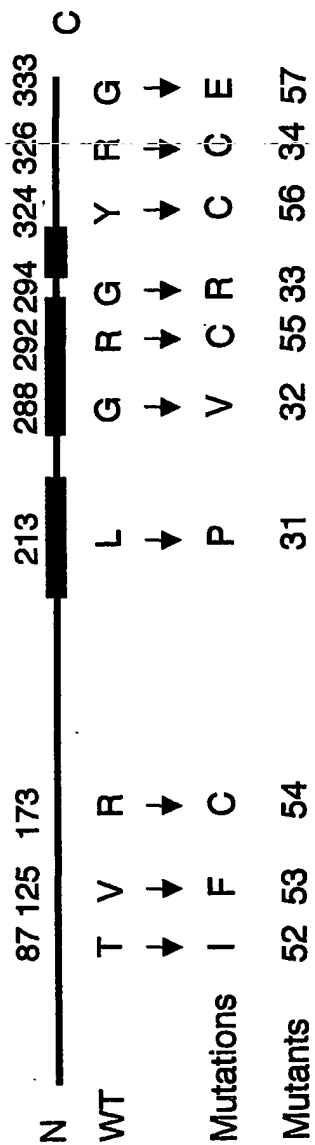
19/40

FIG. 9B



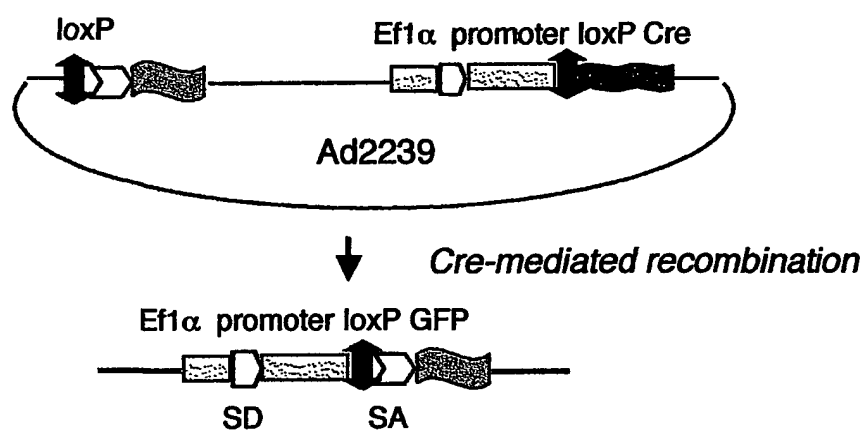
20/40

FIG. 10A



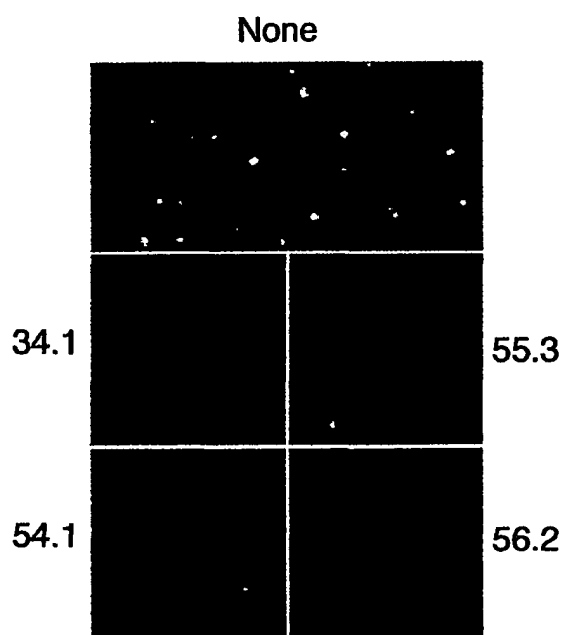
21/40

FIG. 10B



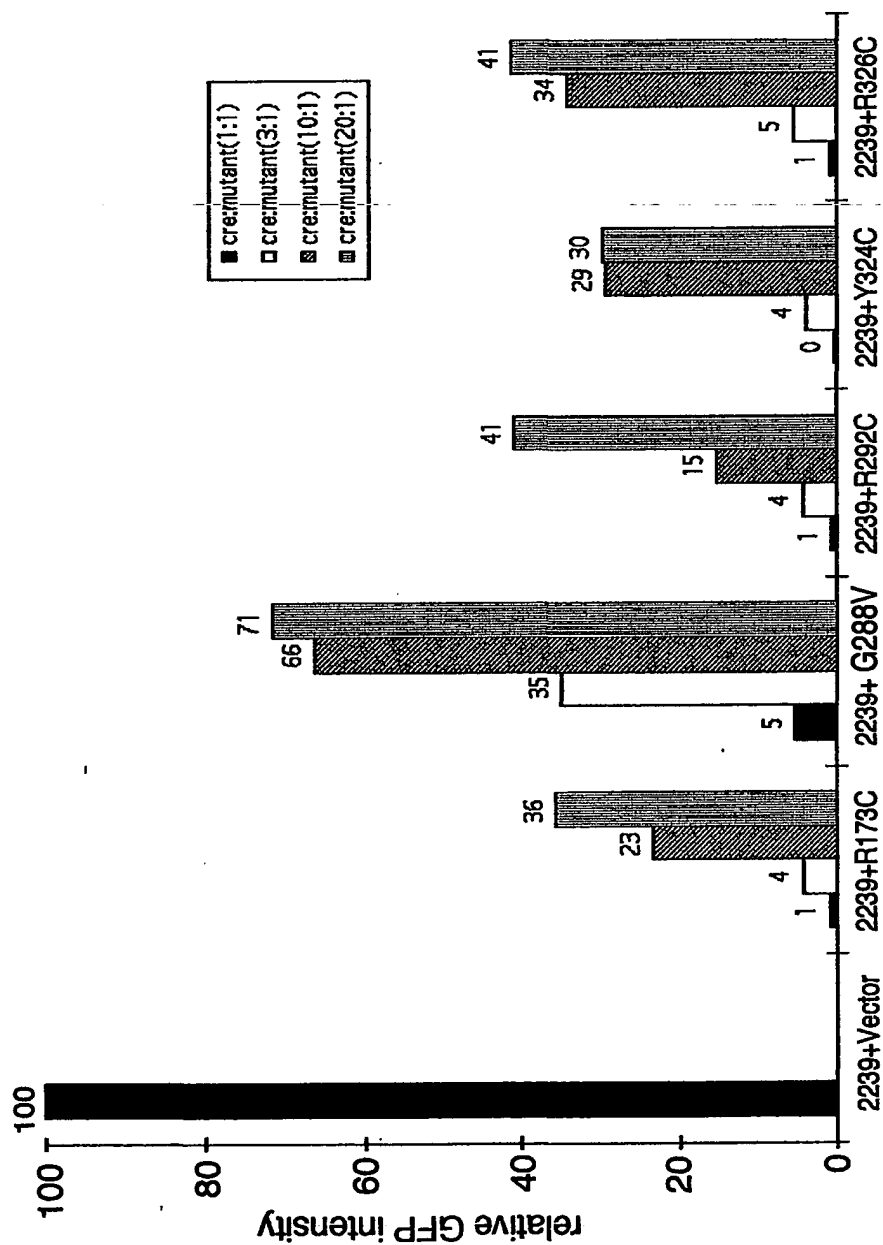
22/40

FIG. 10C



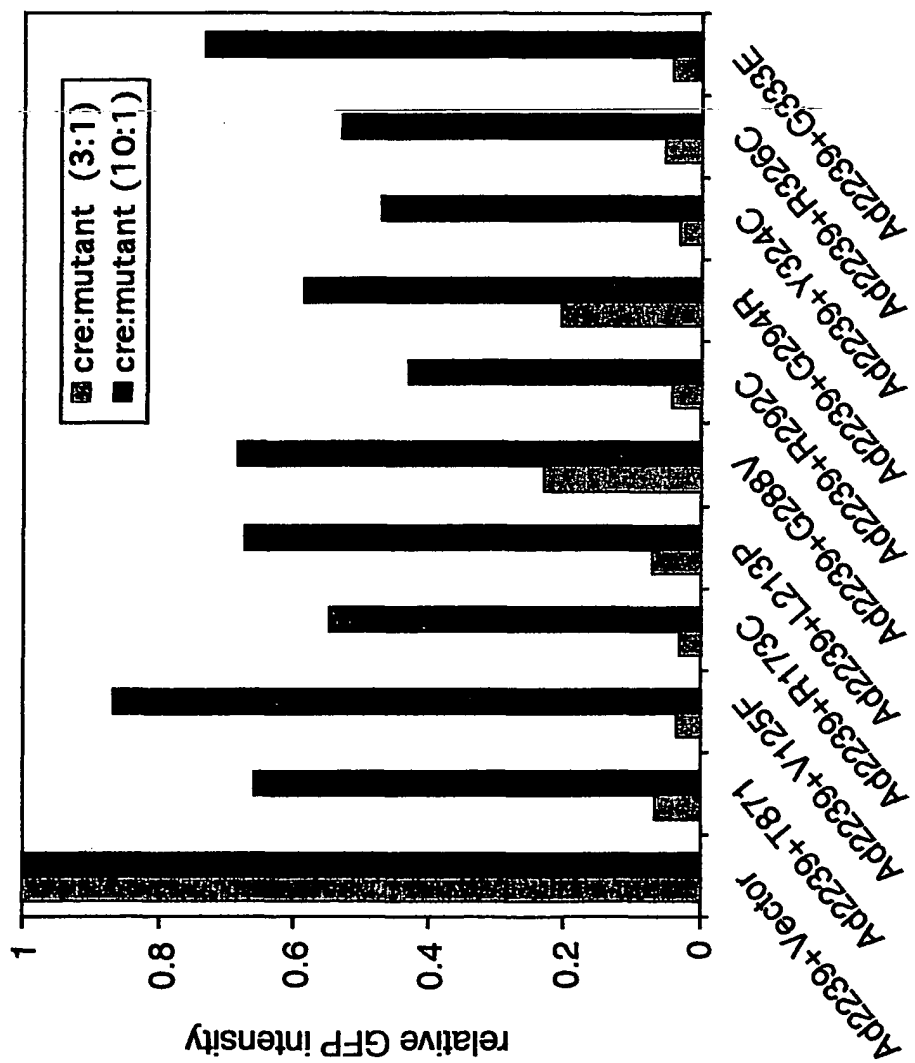
23/40

FIG. 10D



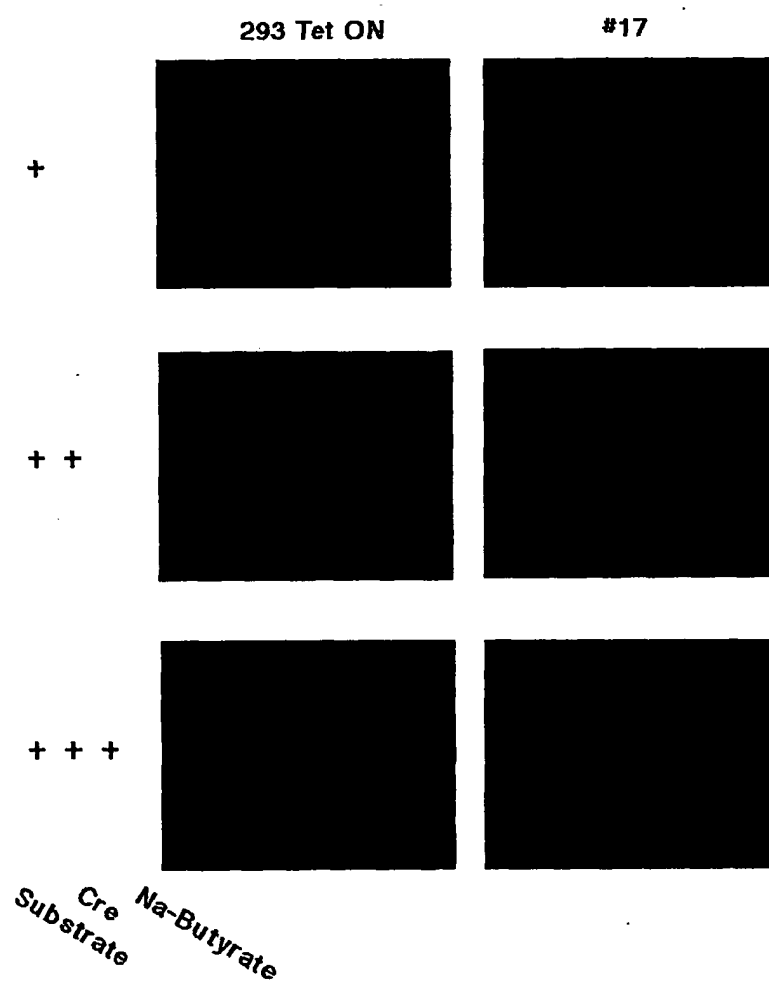
24/40

FIG. 10E



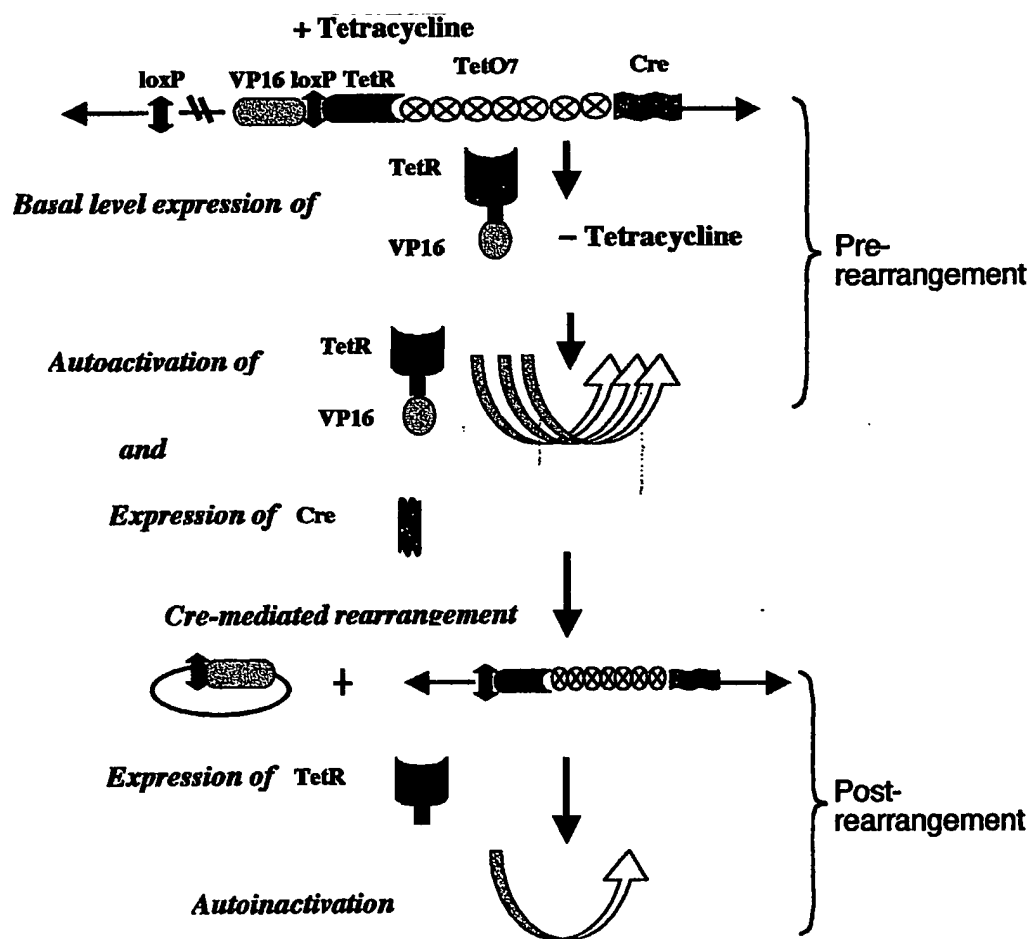
25/40

FIG. 11



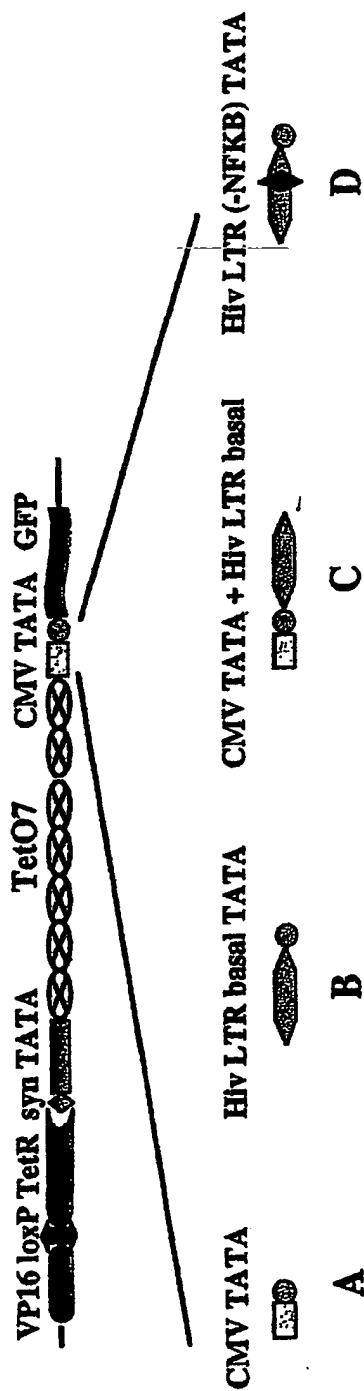
26/40

FIG. 12



27/40

FIG. 13A



28/40

FIG. 13B

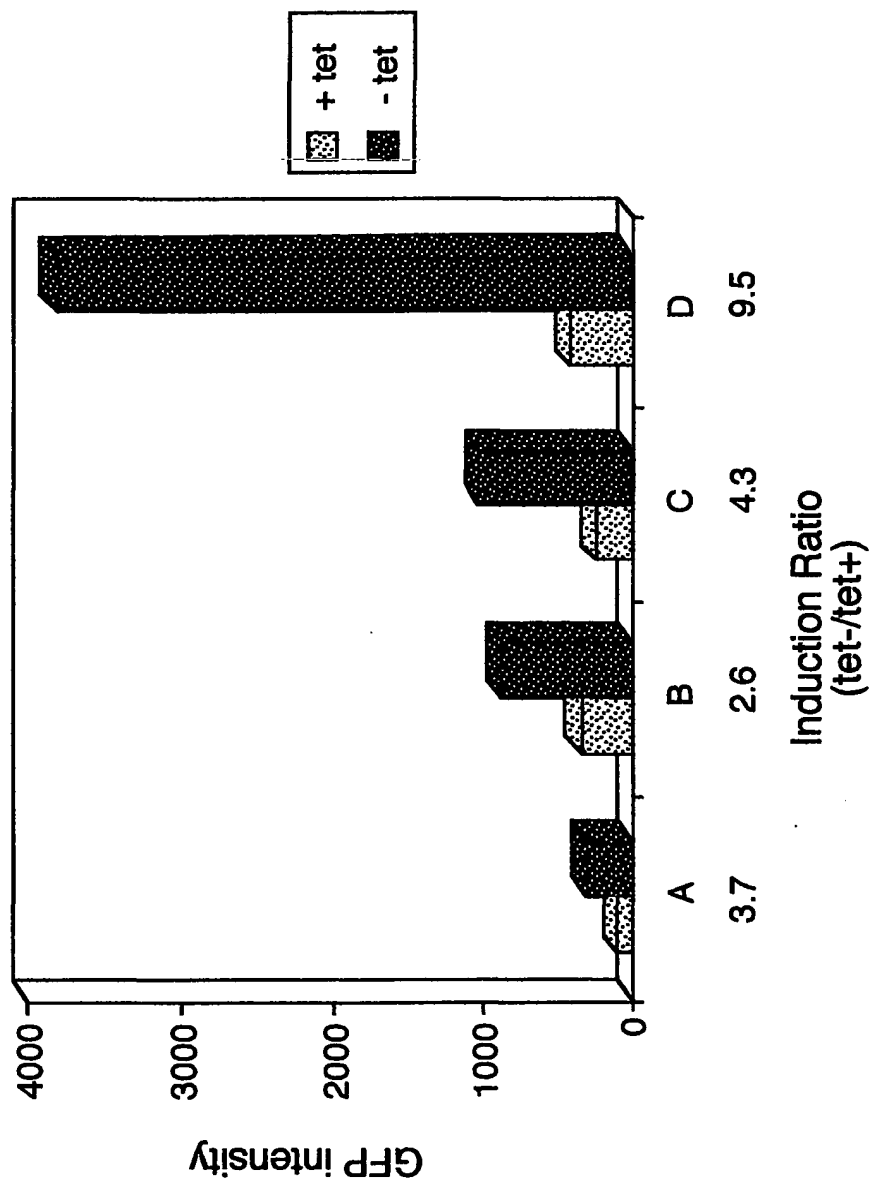
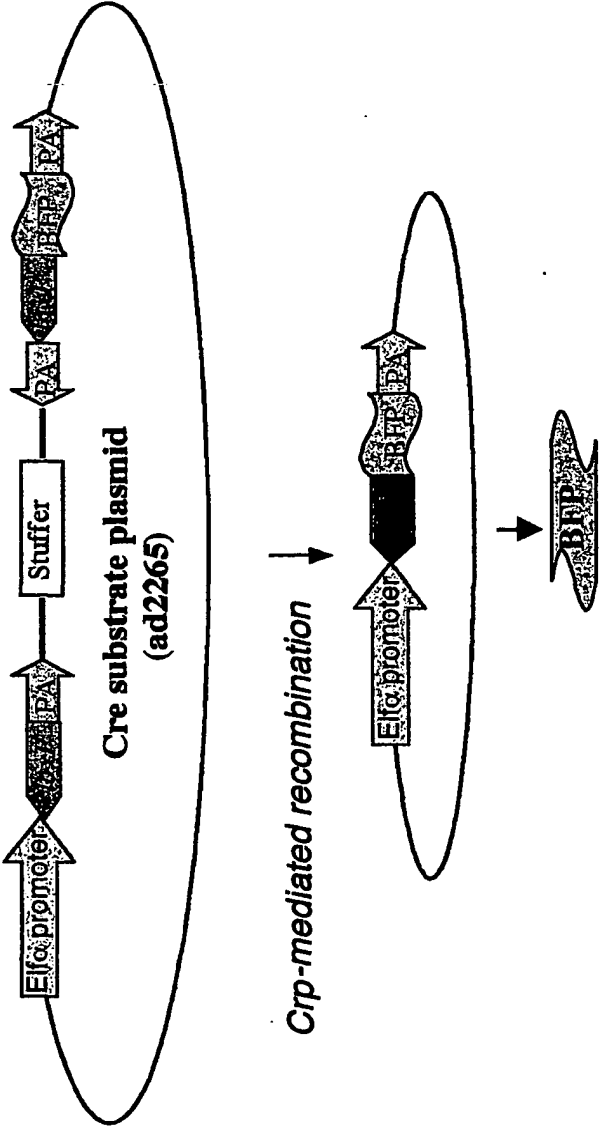


FIG. 14



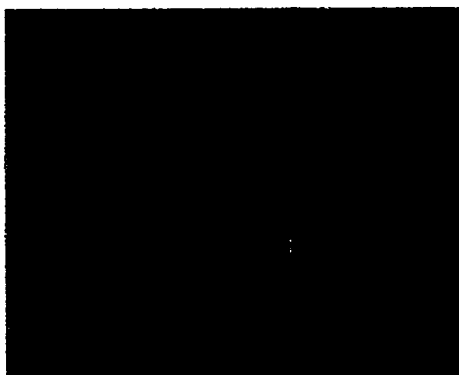
30/40

FIG. 15A

0 μ M Est.

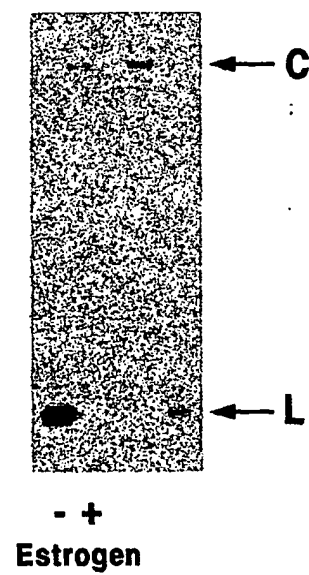


2 μ M Est.



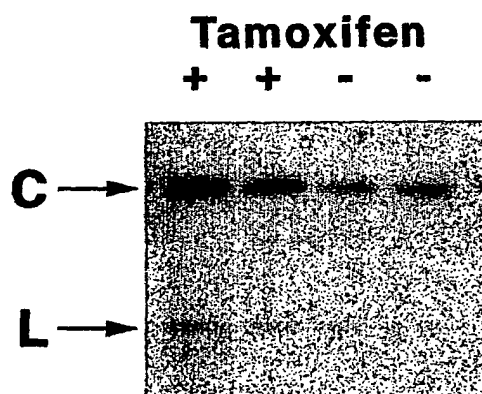
31/40

FIG. 15B



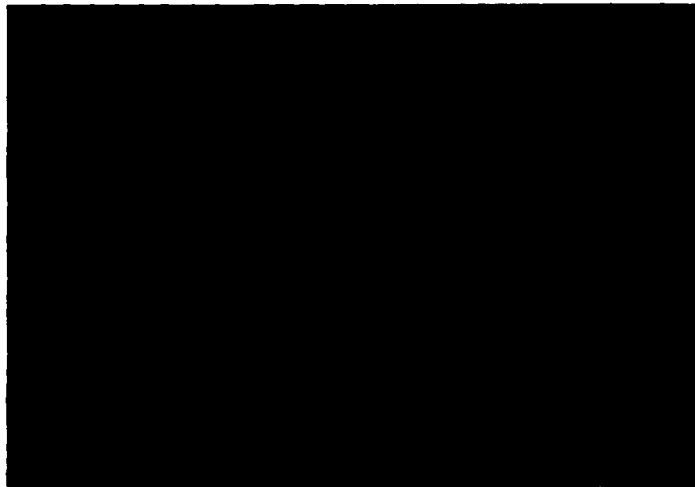
32/40

FIG. 16



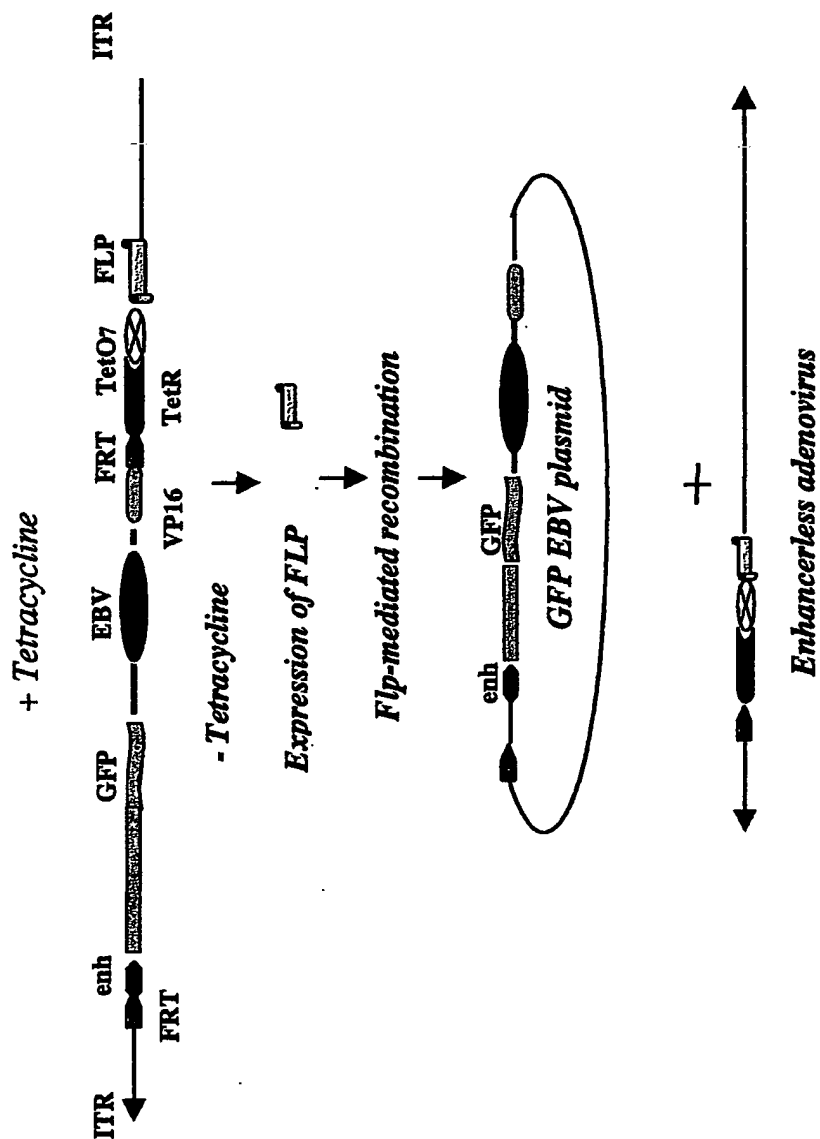
33/40

FIG. 17



34/40

FIG. 18A



35/40

FIG. 18B

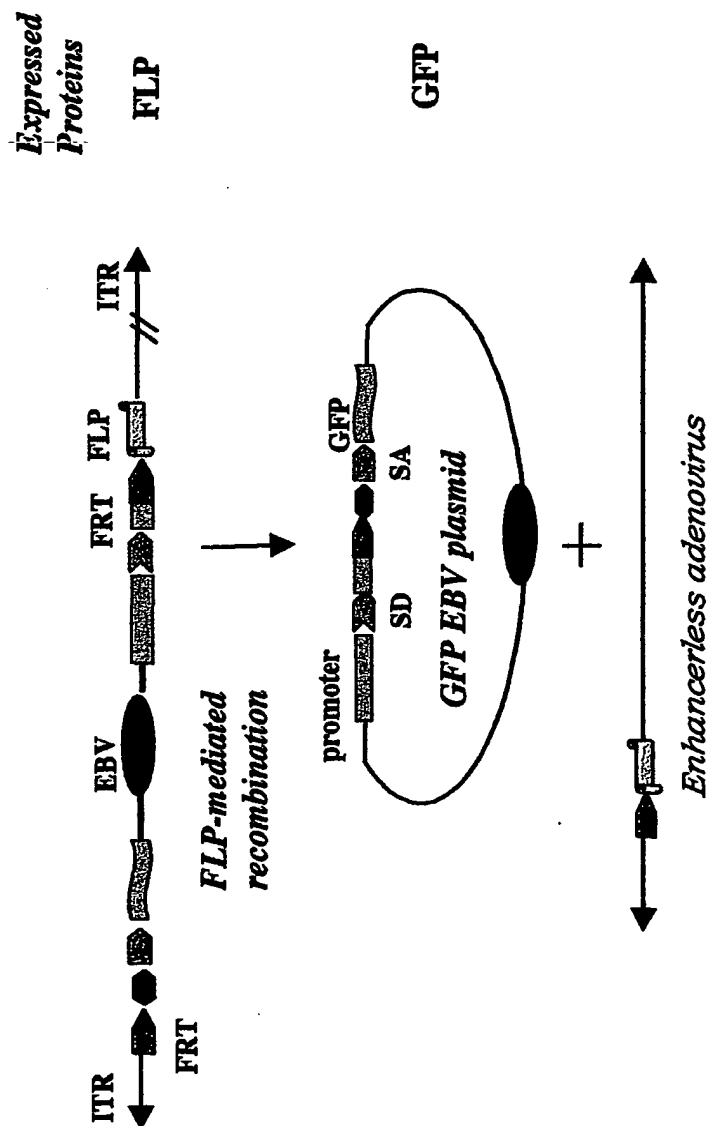
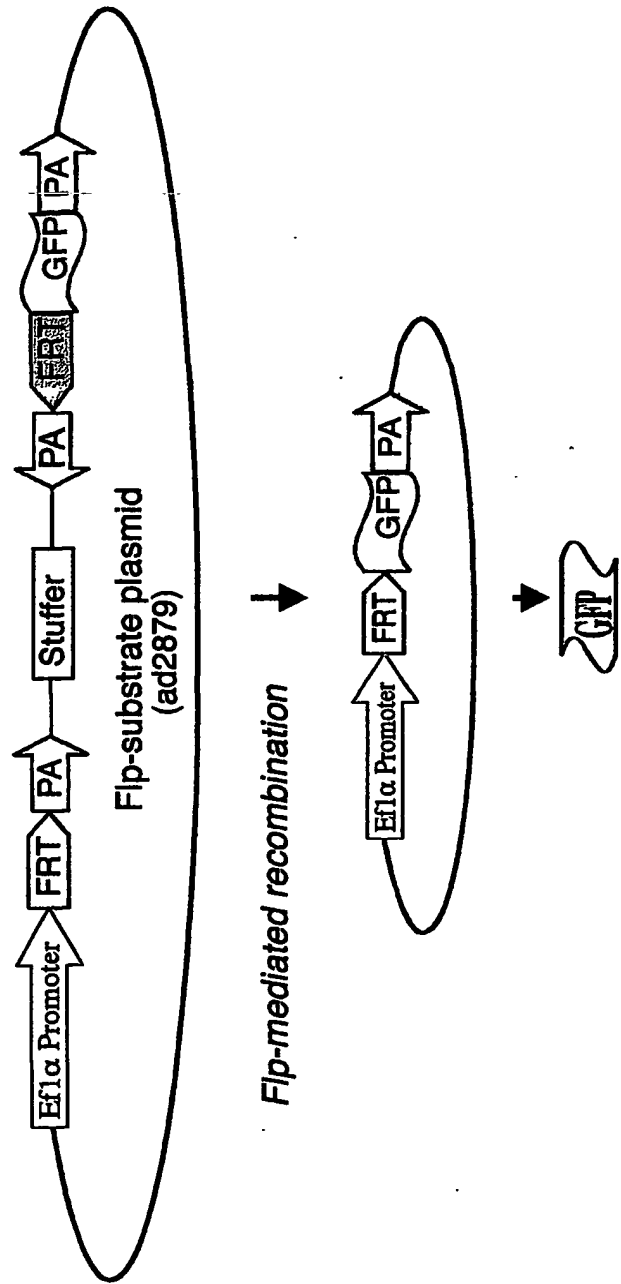
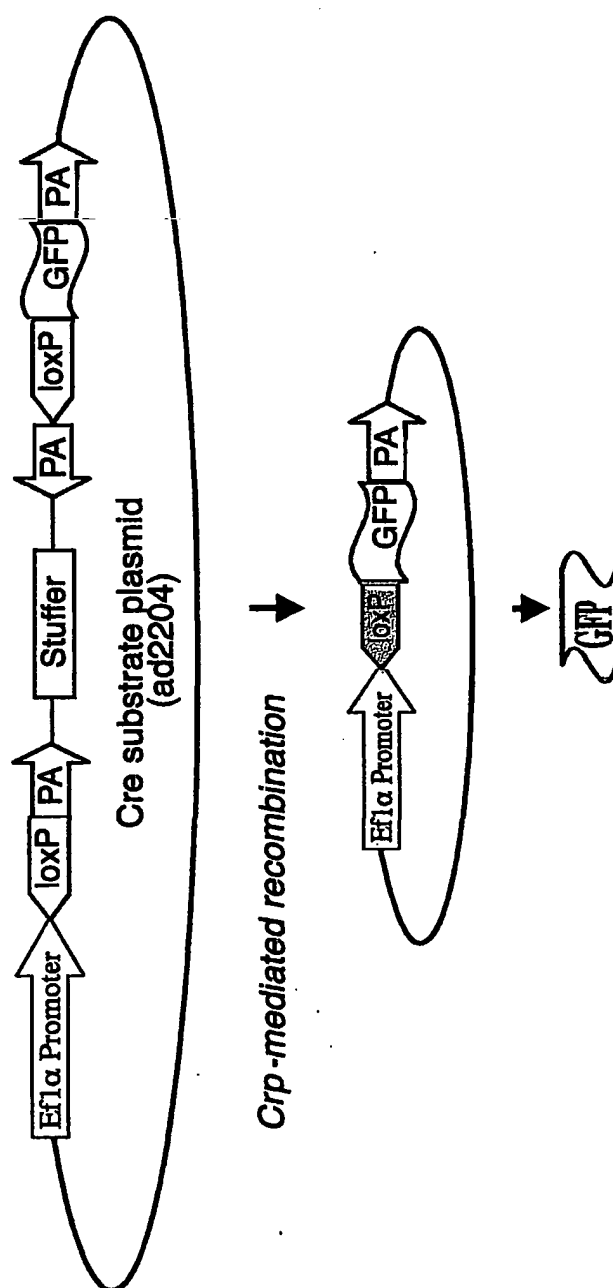


FIG. 19A



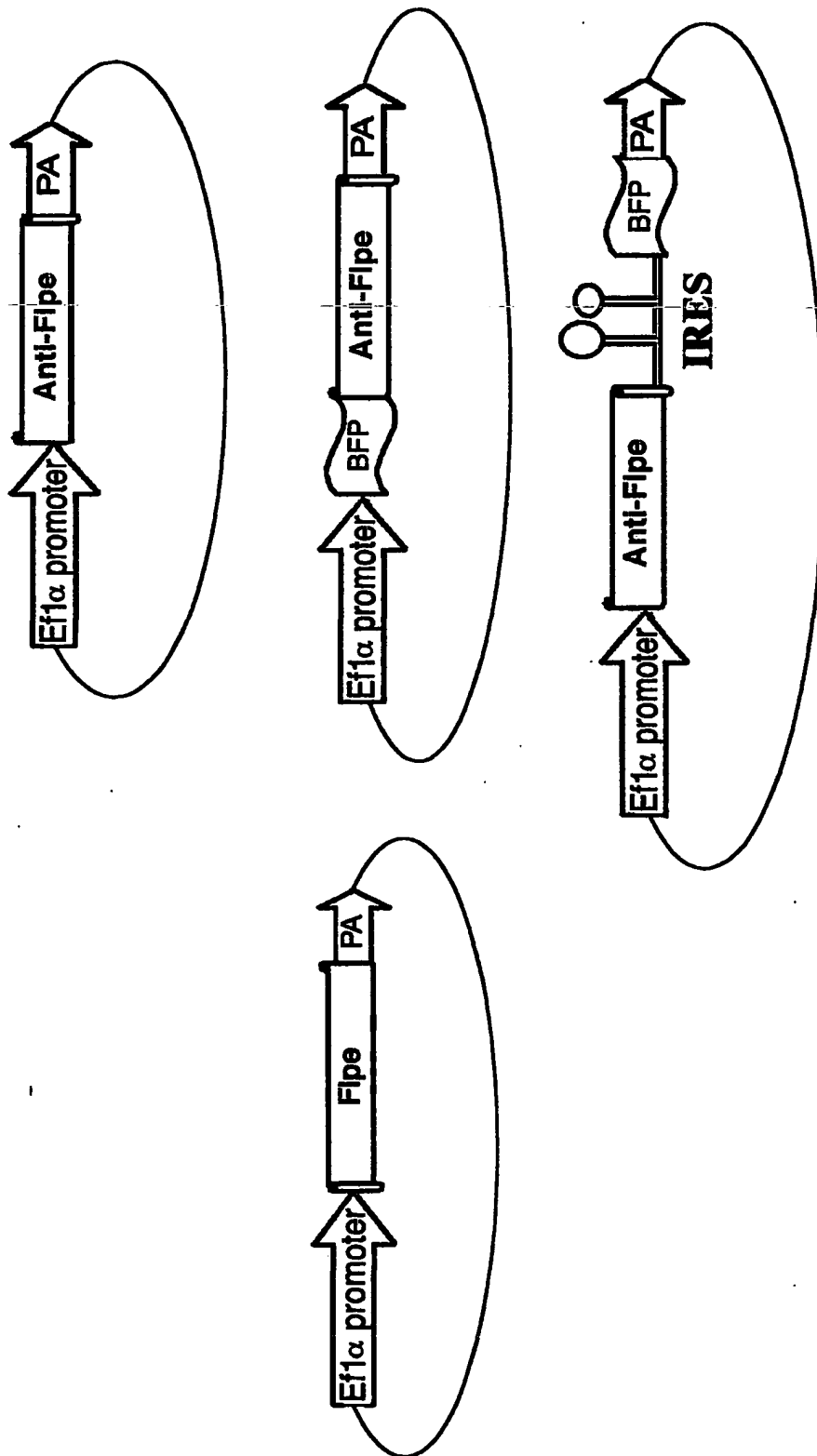
37/40

FIG. 19B



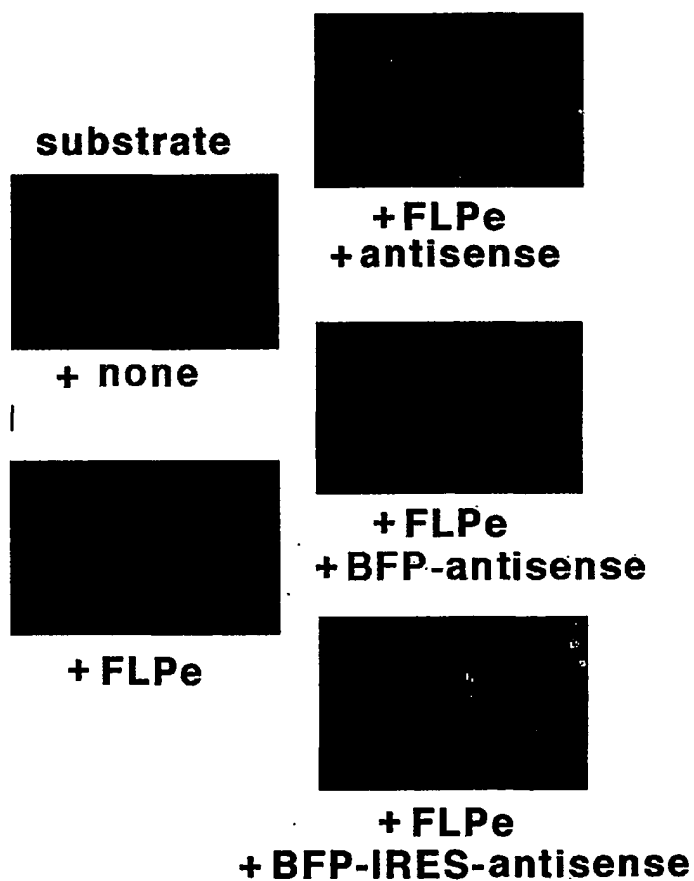
38/40

FIG. 20



39/40

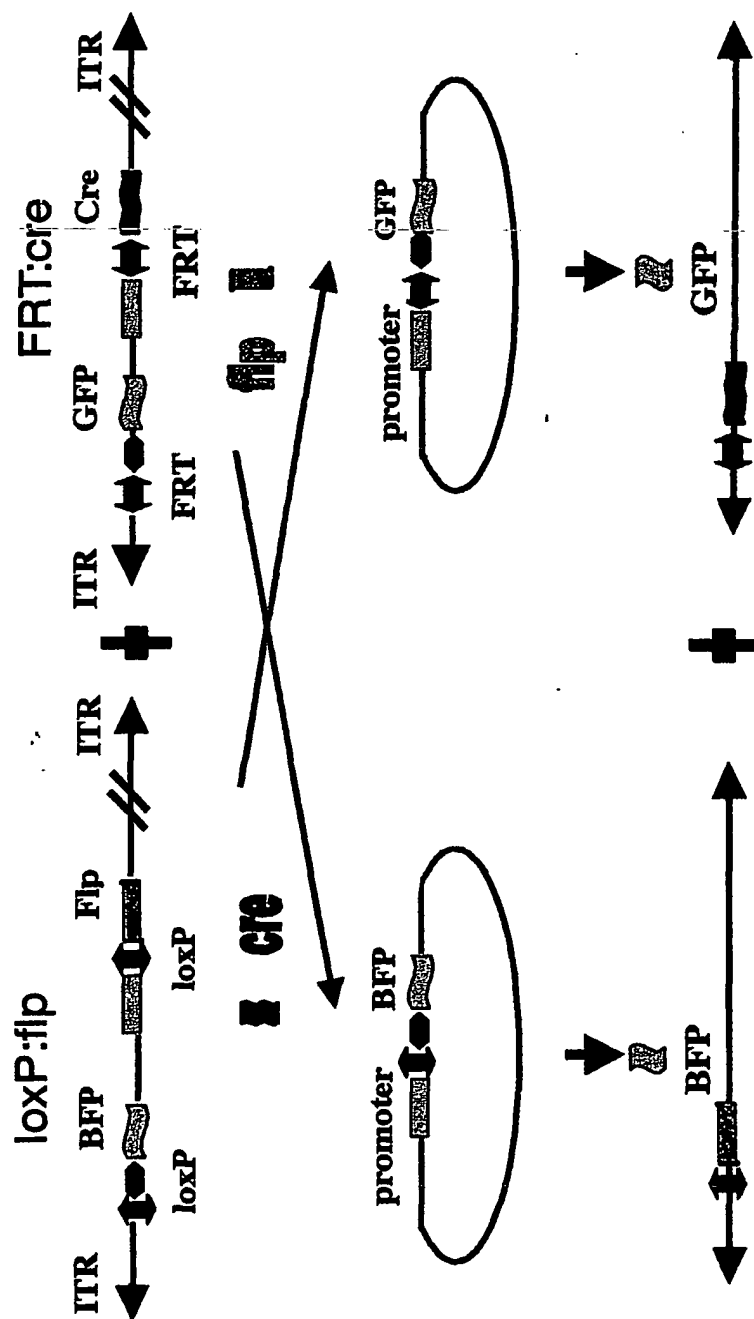
FIG. 21



RO/A

40/40

FIG. 22



SEQUENCE LISTING

<110> The General Hospital Corporation

<120> Self-rearranging DNA vectors

<130> 00786/352W03

<150> US 60/231,053

<151> 2000-09-08

<150> US 60/246,904

<151> 2000-11-08

<160> 10

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2341

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 1

```

gaatcggcca ggcggaattc gattatcatc atcataatat accttatttt ggattgaagc 60
caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcggtggga acggggcggg 120
tgacgtaggt tttagggcgg agtaacttgc atgtattggg aattgtagtt tttttaaaat 180
gggaagttac gtacgcggca tcgatgcgcg ggatatcgcg gcggctagcg acatgaggtt 240
gccccgtatt cagtgtcgct gatttgtatt gtctgaagtt gtttttacgt taagttgatg 300
cagatcaatt aatacgatac ctgcgtcata attgattatt tgacgtggtt tgatggcctc 360
cacgcacggt gtgatatgta gatgataatc attatcactt tacgggtcct ttccggtgat 420
ccgacaggtt acggggcggg gacctcgcgg gttttcgcta tttatgaaaa ttttccggtt 480
taaggcggtt ccgttcttct tcgtcataac ttaatgtttt tttttaaaat accctctgaa 540
aagaaaggaa acgacaggtg ctgaaagcga ggctttttgg cctctgtcgt ttcttttctc 600
tgtttttgtc cgtggaatga acaatggaag ttaacggatc caggccgcga gcaaaaggcc 660
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 720
ccctgacga gcatcacaaa aatcaacgct caagtcagag gtggcgaaac ccgacaggac 780
tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgacct 840
tgccgcttac cggatacctg tccgccttct tcccttcggg aagcgtggcg ctttctcata 900
gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc 960
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 1020
acccggttaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 1080
cgaggtagtg aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 1140
gaagaacagt atttggatc tcgctctgac caaagccagt taccttcgga aaaagagttg 1200
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 1260
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1320
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatcaga ttatcaaaaa 1380
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1440
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1500
tctgtctatt tcgttcatcc atagttgcct gactccccgt agttagata actacgatac 1560
gggagggcgt accatccggc cccagtgcgt caatgatacc gcgtgacca cgctcaccgg 1620
ctcctgattt atcagcaata aaccagccag ccggaagtgc cgagcgaga agtggtcctg 1680
caactttatc cgcctccatc cagtctatta gttgttgccg ggaagctaga gtaagtagtt 1740

```

```

cgccagttaa tagttttcgc aacgttggtg ccattgtctac aggcacgtg gtgtcacgct 1800
cgctggtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 1860
cccccatggt gtgcaaaaaa gcggttagct ccttcgggtcc tccgatagtt gtcagaagta 1920
agttggccgc agtggttatca ctcatggtta tggcagcact gcataattct cttactgtca 1980
tgccatccgt aagatgcttt tctgtgactg gtgagtattc aaccaagaat acgggataat 2040
accgcgccac atagcagaac tttaaaagtg ctcatcattg ggaaacgttc ttcggggcga 2100
aaactctcaa ggaatcttacc gctgttgaga tccagttcga tgtaaccac tcgcgcaccc 2160
aagtgatctt ctgcactctt tactttcacc agcgtttctg ggtgagcaaa aacagggaagg 2220
caaaatgccg caaaaaaggg aataaggggcg acacggaaat gttgaatact catacttttc 2280
ctttttcaat attattgaag catttatcag ggttattgtc tcacacgcgg atacatattt 2340
g

```

<210> 2

<211> 34616

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 2

```

gaatcggcca gcgcgaatta actataacgg tcctaaggta gcgtcatcat cataatatac 60
cttatttttg attgaagcca atatgataat gagggggtgg agtttgtgac gtggcgcggg 120
gcgtgggaac gggcggggtg acgtagggtt tagggcgagg taacttgcac gtattgggaa 180
ttgtagtttt tttaaaatgg gaagttacgt atcgtgggaa aacggaagtg aagatttgag 240
gaagttgtgg gttttttggc tttcgtttct gggcgtaggt tcgcgtgcgg tttctcgggt 300
gtttttttgt gactttaacc gttacgtcat tttttagtc tatatatact cgctctgtac 360
ttggcccttt ttacactgtg actgattgag ctggtgccgt gtcgagtggt gttttttaat 420
agggtttttt actggttaagg ctgactgtta tggctgccgc tgtggaagcg ctgtatgttg 480
ttctggagcg ggagggtgct attttgccca ggcaggaggg tttttcagggt gtttatgtgt 540
ttttctctcc tattaatttt gttataacct ctatgggggc tgtaatgttg tctctacgcc 600
tgcgggtatg tattcccccg ggctatttcg gtcgcttttt agcactgacc gatgttaacc 660
aacctgatgt gtttaccgag tcttacatta tgactccgga catgaccgag gaactgtcgg 720
tggtgctttt taatcacggg gaccagtttt tttacgggta cgcggcgatg gccgtagtc 780
gtcttatgct tataagggtt gtttttcctg ttgtaagaca ggcttctaata gtttaaatgt 840
ttttttttgt tatttttttt tgtgtttaat gcaggaaacc gcagacatgt ttgagagaaa 900
aatgggtgtc tttctgtggg tgggtccgga acctacctgc ctttatctgc atgagcatga 960
ctacgatgtg cttgcttttt tgcgcgaggg tttgcctgat tttttgagca gcaccttgca 1020
ttttatatcg ccgcccatgc aacaagctta cataggggct acgctggtta gcatagctcc 1080
gagtatgcgt gtcataatca gtgtgggttc ttttgtcatg gttcctggcg gggaaagtggc 1140
cgcgtgggtc cgtgcagacc tgcacgatta ttttcagctg gccctgcgaa gggacctacg 1200
ggatcgcggg atttttggtt atgttccgct tttgaatctt atacaggtct gtgaggaacc 1260
tgaatttttg caatcatgat tcgctgcttg aggtgaagg tggagggcgc tctggagcag 1320
atttttaca tggccgggact taatattcgg gatttgctta gagacatatt gataagggtg 1380
cgagatgaaa attatttggg catggttgaa ggtgctggaa tgtttataga ggagattcac 1440
cctgaagggt tttagccttta cgtccacttg gacgtgagg cagtttgctt tttggaagcc 1500
attgtgcaac atcttacaac tgccattatc tgttctttgg ctgtagagtt tgaccacgcc 1560
accggagggg agcgcgttca cttaatagat cttcattttg aggttttgga taatcttttg 1620
gaataaaaaa aaaaaaaca tggttcttcc agctcttccc gctcctcccg tgtgtgactc 1680
gcagaacgaa tgtgtaggtt ggctgggtgt ggcttattct gcggtggtgg atgttatcag 1740
ggcagcggcg catgaaggag tttacataga acccgaagcc agggggcgcc tggatgcttt 1800
gagagagtgg atatactaca actactacac agagcgagct aagcgacgag accggagacg 1860
cagatctggt tgtcacgccc gcacctggtt ttgcttcagg aaatatgact acgtccggcg 1920
ttccatttgg catgacacta cgaccaacac gatctcgggt gtctcggcgc actccgtaca 1980
gtaggggatcg cctacctcct tttgagacag agaccgcgc taccatactg gaggatcatc 2040
cgctgctgcc cgaatgtaac actttgacaa tgcacaacgt gagttacgtg cgaggtcttc 2100
ctgcagtggt gggatttacg ctgattcagg aatgggttgt tccctgggat atgggtctga 2160
cgcgggaggga gcttgaatac ctgaggaagt gtatgcacgt gtgcctgtgt tgtgccaaca 2220

```

ttgatatcat	gacgagcatg	atgatccatg	gttacgagtc	ctgggctctc	cactgtcatt	2280
gttccagtc	cggttccctg	cagtgcatag	ccggcgggca	ggttttggcc	agctgggtta	2340
ggatgggtgt	ggatggcgcc	atgtttaatc	agaggtttat	atggtaccgg	gaggtggtga	2400
attacaacat	gccaaaagag	gtaatgttta	tgtccagcgt	gtttatgagg	ggctgccact	2460
taatctacct	gcgcttgagg	tatgatggcc	acgtgggttc	tgtgggtccc	gccatgagct	2520
ttggatacag	cgccttgac	tgtgggattt	tgaacaatat	tgtgggtgctg	tgctgcagtt	2580
actgtgctga	tttaagttag	atcagggtgc	gctgctgtgc	ccggaggaca	aggcgtctca	2640
tgctgcgggc	ggtgcgaatc	atcgctgagg	agaccactgc	catgttgtat	tcctgcagga	2700
cggagcggcg	gcggcagcag	tttattcgcg	cgctgctgca	gcaccaccgc	cctatcctga	2760
tgacagatta	tgactctacc	cccatgtagg	cgtggacttc	cccttcgccg	cccgttgagc	2820
aaccgcaagt	tgacagcag	cctgtggctc	agcagctgga	cagcgacatg	aacttaagcg	2880
agctgcccgg	ggagtttatt	aatatcactg	atgagcgttt	ggctcgacag	gaaaccgtgt	2940
ggaatataac	acctaagaat	atgtctgtta	cccatgatata	gatgcttttt	aaggccagcc	3000
ggggagaaag	gactgtgtac	tctgtgtgtt	gggaggagg	tggcaggttg	aatactaggg	3060
ttctgtgagt	ttgattaaag	tacggtgatc	aataaagct	atgtgggtgt	ggggctatac	3120
tactgaatga	aaaatgactt	gaaattttct	gcaattgaaa	aataaacacg	ttgaaacata	3180
acatgcaaca	ggttcacgat	tctttattcc	tgggcaatgt	aggagaaggt	gtaagagttg	3240
gtagcaaaag	tttcagtggt	gtattttcca	ctttcccagg	accatgtaaa	agacatagag	3300
taagtgttta	cctcgctagt	ttctgtggat	tcactagtgc	cattaagtgt	aatggtaagt	3360
atcataggtt	tagttttatc	accatgcaag	taaacttgac	tgacaatgtt	atttttagca	3420
gtttgacttt	gggttttttg	ataggctaga	aggttaggca	taaatacaac	tgcatattgt	3480
tatggatttg	cattagttga	gttcccattt	ctaaagtctc	agtaatgttt	tttaagttag	3540
gagttctcca	ttagaacacc	gttttggtca	aatactaagg	ataactaac	acttgcaacg	3600
gtgcctgtca	tggatgaaag	atctccagat	acagccaaag	cagctacagt	agctagtact	3660
tgactcccac	attttgtaag	aaccaaagta	aatttgcaat	cattatctga	atgaattctg	3720
cagttaggag	atgggtctgg	ggttgctcac	agggtaaagt	tgctcatcatt	tttgtttctc	3780
attgtaatgg	cccctgagtt	gtcaaaagct	aaacccgctc	caagttagt	aatcatggca	3840
ccgttttcat	tgtaatcaat	gccagagcca	attttagttt	ttattgggtt	gatatctgga	3900
gactcagatg	tgtttgatc	aaactccaga	ccctttcctg	catttatagc	tatggcagta	3960
ttatcaaagt	ttagtccact	ggattttttt	atgctaactt	ccagtttttt	agtattgttt	4020
gatgcattaa	aaaggtatag	gcctctgtta	tagttttatg	ccaagttatg	agatgcatta	4080
atatacagg	gtccctgccc	cagtttaaga	cgtagttttg	tttgagcatc	aaatgggtaa	4140
ttccatctca	gaattaacaa	gttggtattt	atacgcatgc	caccgcccgt	tttaatttcc	4200
atgttggttg	atgaatcata	accaatagct	cctgcaactt	tggttctaag	ggagttttgt	4260
tcaacgggtga	cacctggtcc	agtaactact	gttagtgtat	cggagtgttg	tgctacttgc	4320
aaaggaccgc	ttattttaat	tcctattttt	ccattattta	cataaatagg	atcttccatt	4380
ttaagtggct	gagtaacagt	ggttacattt	agcgggggtg	atgcagttac	agtaaggggtg	4440
tcgctgtcac	tgccagagag	gggggctgat	gtttgcaggg	ctagctttcc	atctgacact	4500
gtaatgggccc	ctttagtagc	aatgcttagt	ttggagtctt	gcacggtcag	tggggcttgt	4560
gactgtacgc	taagagcgcc	gctagtaact	atcagaggag	cgggtggttg	cactgttagg	4620
gcgcctgagg	taattgtaag	tgggtgcggag	gtgtccaaac	ttatgtttga	ctttgttttt	4680
tttaagtggct	gagtaacagt	ggttacattt	tgggaggtga	ggtttccggc	cttgtctagg	4740
gtaagaccgc	tgcccatttt	aagcgcaagc	atgccgtggg	aggtgtccaa	aggttccggag	4800
acgcgtagag	agagaactcc	agggggactt	tcttggaac	cattgggtga	aacaaatgga	4860
ggggtaaaga	agggcacagt	tggaggcccg	gtttctgtgt	catatggata	cacgggggtg	4920
aaggtgtctt	cagacggtct	ggcgcttttc	atctgcaaca	atatgaagat	agtgggtgag	4980
gagggacaag	aacatgagga	atgttgacatc	ccatttaaac	tttgagagaa	gtttgcagct	5040
aaaaggcggc	tgagatacca	gagttgggag	gaaggaaagg	aggtgatgct	gaataagctg	5100
gacaaagatt	tgctgactga	ttttaagtaa	gtaattttatt	gtgtgtttat	gttagttgaa	5160
tggaataaga	tctctaatac	cacacatggt	tttaataaga	gtgcagaggt	cctctggacc	5220
ctgatagggg	aagtgcaggc	agccctctgt	ttctgcccag	tgctgggtga	cgggtgatag	5280
tttttctccc	accataagca	ccagtttttg	gcgctgggtg	ggtagcttgt	agctgaggcg	5340
gttgccggta	gtgggttttt	cgtaggtaag	tttggcctgc	ttgaccacac	aaaagatacc	5400
tcttttacac	tgggttaggt	taaccatgtc	ttcaacttct	tgtttttaggc	gttctcgctc	5460
ggacgcggcc	ttgcgccttt	ctagtaggcg	ctgttcgggtg	tttaattccat	ccaattctag	5520
atctagagat	tcagtcatct	ccacctgtca	aattaaagta	gctaattctca	gtgggggtgg	5580
gagaaggggg	gcgaggctga	ttgattgggg	caataacctg	ttgcagtggt	atgacagcgg	5640
gcactgggaa	agtaggggtg	ttcatggcat	ctatggcatt	ccagccaatg	tcaaggtatg	5700

gatatatggc	tagggcaaaa	atggtactgc	aaaaaacccat	gacagagatg	atggcggtata	5760
accaggtctc	tgacaaatcg	ctctgtttgt	tgtagcagct	gggaatgttc	catattttgag	5820
tgaatctgca	ggaaatatgt	cttttgggag	gcgctgaggt	ttgggagcaa	agcacaggta	5880
ggcgcaaaa	aatcagcaaa	acaaaaatga	cactccgttt	cataattaaa	gaattctgag	5940
aagatcagct	atagtcctgt	ctctgtattg	cggatggtgc	ctgaggtacg	caatgcgcac	6000
acaaaccag	tcaatgaact	gaatgaaggc	gatgactaca	gtgacgaggc	tgcagatgag	6060
gataagggtg	acaaatccgt	aaagcaggta	aactgtgaaa	gggtgggatgc	aatctacttc	6120
gatgtgagcg	accgcggcca	atgtagagca	cgcacagaaa	agcgcaacaa	gggtcaataa	6180
tataagaact	cgaggaatca	tgtctcattt	aatcatactg	taaaagaaga	gaacatggtt	6240
tcagaccgtc	caatctatga	attttttcat	tgtgtgggtt	gagcacaatg	ataggcctat	6300
agatgggggg	tctggcgctg	ctgcgcttta	ggcaacaaat	aagccacata	ataataaggc	6360
aaacaaacat	aagcgctatg	gaaaaccacc	acatgtccaa	gctcgcccag	tcattgacaa	6420
aggcatgaac	ttggggtaaa	tttagggcag	atgttagtcc	ggtagcagtg	gtgttgcgat	6480
agtccgttgt	gggcgcgatg	gttagccggg	tcactctctg	agcaggcaag	ctgaagctgg	6540
gtttgatcaa	atttgcagtg	caggcgctgg	cagaaatcag	gcgctaacgt	ccaggaaagt	6600
ttgatttgaa	ggttgtgggt	ataatcttgc	ccgcctggag	catatcccac	atagagttaa	6660
ttgtccaggg	gaatacaagc	aagcggaaaa	tcaaggcatt	ttcttttcat	caataaaact	6720
gcgtctgctt	ttgtatttga	gataaagtaa	ggtacatacc	aaagcaagcg	ctgtaataag	6780
cagagcgggtg	gaacaaaagg	tgccagtgtt	ctctaaacac	ttttgtgggg	gccacaactt	6840
gtactgtttg	ctcatgtaca	tggtaatatc	gcacatttca	taaaatggaa	atttatacat	6900
aaaagtttta	cgattttcac	cttggaagac	tgtgacatta	tagtcgttag	tgtcacctag	6960
ctgccaaaata	gcataacag	cataactgcc	aattttgtct	ttgtggcgaa	taataagctt	7020
ttcatgttct	gtggtgcatt	ttataagagt	agtgcattca	ttagcttctg	atttaaagt	7080
aacattgcaa	gctggttcct	taaactcaac	ctttttggca	gcgctgcaga	ctgccgcaag	7140
ggcgagcaag	cctaaaatca	tgtacctcat	cttggtgtgt	gccccagcg	tttaaaaagc	7200
tgacaatagg	tacaaacgtg	cgtgcagcag	cgggcaacc	taaggcacag	aagtgcctag	7260
ataagaataa	acagaattac	aagagttaag	ataacccga	ccccaattcc	agaaaaatta	7320
gacaagcttg	tagagttact	tgaattgctc	atatacttaa	ttaaaaaatc	ccagcacccc	7380
gcaaaatgct	tttttgacct	gagttccggg	agttgagctc	acctcctggt	ttggaaaaat	7440
gggagtaatg	tctggttacg	ctcaggctgt	aggtgtgggc	gcagcaaccg	gtgacgcact	7500
cgtacgttcc	cggcagggtga	ggagggtggt	ggtggtggtg	tttttcttga	cgggtgagtt	7560
gaagccgaga	aggttgtgtg	gcaaaacttac	ttcgtctcgc	tggaaactgt	tgtaaattac	7620
aaatgaagag	ccgttaaagt	accaggttaag	gtacttattg	gcccgttgt	gcaaaccgga	7680
ggtgaggttt	gctttggtct	gctttgggtg	ggtaaaaacg	gtggcggttca	caggatggcg	7740
acaggagccc	cagttagattc	taatttctgt	atttattata	ctcagcacag	agatgacaac	7800
aaagatcttg	atgtaatcca	gggttaggac	atgtgcaaac	cacggtcaga	acacagggac	7860
ccgctccccg	ctccactagc	agggggcgct	tggtaaactc	ccgaatcagg	ctacgtgtaa	7920
gctctacctg	ggtggtgagc	cggacgccgt	gcgcggggcc	ctcgatatgc	tcttcggggca	7980
attcaaagta	acaaaactca	ccggagccgc	gggcaaaagca	cttgtggcgg	cggcagtggt	8040
cgagggtgtg	caggcgcagt	cgctctgcct	ctccactggt	cattcagtcg	tagccgtccg	8100
ccgagtcctt	caccgcgtca	aagttgggaa	taaactggtc	cgggtagtg	ccgggaggtc	8160
cagaaaaggg	gttgaagtaa	accgaaggca	gaaactcctc	aataaattgt	agagttccaa	8220
tgcctccgga	gcgcggctcc	gaggacgagg	tctgcagagt	taggatcgcc	tgacggggcg	8280
taaatgaaga	gcggccagcg	ccgccgatct	gaaatgtccc	gtccggacgg	agaccaagag	8340
aggagctcac	cgactcgtcg	ttgagctgaa	tacctcgccc	tctgattttc	aggtgagtta	8400
tacctgccc	ggcgaccgc	accctgtgac	gaaagccgc	cgcaagctgc	gcccctgagt	8460
tagtcatctg	aacttcggcc	tggcgctctc	tgggaagtac	cacagtgggtg	ggagcgggac	8520
tttctgtgta	caccagggca	gcggggccaac	tacggggatt	aaggttatta	cgaggtgtgg	8580
tggtaatagc	cgcctgttcg	aggagaattc	ggtttcgggtg	ggcgcggtatt	ccgttgaccc	8640
gggatatcat	gtgggttccc	gcgctcatgt	agtttattcg	ggttgagtag	tcttgggcag	8700
ctccagccgc	aagtcaccat	tgtggctggt	aactccacat	gtaggcggtg	ggaatttcct	8760
tgctcataat	ggcgctgacg	acaggtgctg	gcgcgggtg	tggccgctgg	agatgacgta	8820
gttttcgcgc	ttaaatttga	gaaagggcgc	gaaactagtc	cttaagagtc	agcgcgcagt	8880
atattgctgaa	gagagcctcc	gcgtcttcca	gcgtgcgcgc	aagctgatct	tcgcttttgt	8940
gatacaggga	gctgcgggtg	agggagcgca	gagacctgtt	ttttattttc	agctcttgtt	9000
cttggccctt	gctttgttga	aatatagcat	acagagtggg	aaaaatccta	tttctaagct	9060
cgcggtcgga	tacgggttcg	ttgggcgcca	gacgcagcgc	tcctcctcct	gctgctgccg	9120
ccgctgtgga	tttcttgggc	tttgtcagag	tcttgcctatc	cggctgcctt	tgttcttgtg	9180

tgaccgctgc	tggtgctgcc	gctgccgctg	cgcgcggtgc	agtaggggct	gtagagatga	9240
cggtagtaat	gcaggatggt	acgggggaag	gccacgccgt	gatggtagag	aagaaagcgg	9300
cgggcggaag	agatgttgcc	cccacagtct	tgcaagcaag	caactatggc	gttcttgtgc	9360
ccgcgccacg	agcggtagcc	ttggcgctgt	tggtgtctct	gggctaaccg	cgggcgctgc	9420
ttagacttac	cggccctggt	tccagtgggt	tccatctac	ggttgggtcg	gcgaacaggc	9480
agtggcggcg	gcgcctgagg	agcggaggtt	gtagcgatgc	tgggaacggt	tgccaatttc	9540
tggggcgcgc	gcgaggggaa	tgcgaccgag	ggtgacggtg	tttcgtctga	cacctcttcg	9600
gcctcggaag	cttcgtctag	gctgtcccag	tcttccatca	tctcctcctc	ctcgtccaaa	9660
acctcctctg	cctgactgtc	ccagtattcc	tctcgtcccg	tgggtggcgg	cgggcgcgagc	9720
tgcagcttct	ttttgggtgc	catcctggga	agcaagggcc	cgcggctgct	gatagggtcg	9780
cggcgggcgg	gggattgggt	tgagctcctc	gccggactgg	gggtccaggt	aaaccccccg	9840
tccctttcgt	agcagaaact	cttggcgggc	tttgttgatg	gcttgcaatt	ggccaaggat	9900
gtggcccttg	gtaatgacgc	aggcggttaag	ctccgcattt	ggcgggcggg	attggtcttc	9960
gtagaaccta	atctcgtggg	cgtggtagtc	ctcaggtaca	aatttgcgaa	ggtaagccga	10020
ctccacacgc	cccggagtga	gtttcaaccc	cggagcccg	gacttttcgt	caggcgaggg	10080
accctgcagc	tcaaaaggta	cgataatttg	actttcgcta	agcagttgcg	aattgcagac	10140
cagggagcgg	tgcggggtgc	ataggttgca	gcgacagtga	cactccagta	ggccgtcacc	10200
gctcacgtct	tccatgatgt	cggagtggta	ggcaaggtag	ttggctagct	gcagaaggta	10260
gcagtgaccc	caaagcggcg	gagggcattc	acggtactta	atgggcacaa	agtcgctagg	10320
aagcgcacag	caggtggcgg	gcagaattcc	tgaacgctct	aggataaagt	tcctaaagt	10380
ttgcaacatg	ctttgactgg	tgaagtctgg	cagaccctgt	tgagggttt	taagcaggcg	10440
ttcggggaag	ccaggtgcgc	gcgcacggag	cgctcgttga	agggcgctca	agggcgctca	10500
taggtccttc	aagttttgct	ttagcagctt	ctgcagctcc	tttaggttgc	gctcctccag	10560
gcattgtctg	cacacgcca	tggccgtttg	ccaggtgtag	cacagaaata	agtaaacgca	10620
gtcgcggacg	tagtcgcggc	gcgcctcgcc	cttgagcgtg	gaatgaagca	cgttttgccc	10680
gaggcggttt	tcgtgcaaaa	ttccaaggta	ggagaccagg	ttgcagagct	ccacgttgga	10740
aattttgcag	gcctggcgca	cgtagccctg	gcgaaagggtg	tagtgcaacg	tttctcttag	10800
cttgcgctgc	atctccgggt	cagcaaagaa	ccgctgcatg	cactcaagct	ccacggtaac	10860
aagcactgcg	gccatcatta	gcttgcgctg	ctcctccaag	tcggcaggct	cgcgcgtctc	10920
aagccagcgc	gccagctgct	catcgccaac	tgccggtagg	ccctcctcgg	tttgttcttg	10980
caagtttgca	tcctctccca	ggggtcgtgc	acggcgacag	atcagctcgc	tcagtactgt	11040
gctcataacc	ttggggggtg	ggttaagtgc	cgggtaggca	aagtgggtga	cctcgatgct	11100
gcgtttcagc	acggctaggc	gcgcgttgc	accctcaagt	tccaccagca	ctccacagtg	11160
actttcattt	tcgctgtttt	cttggtgcag	agcgtttgcc	gcgcgtttct	cgtcgcgtcc	11220
aagaccctca	aagatttttg	gcacttcgtc	gagcgaggcg	atatcaggta	tgacagcgcc	11280
ctgccgcaag	gccagctgct	tgtccgctcg	gctgcggttg	gcacggcagg	ataggggtat	11340
cttgagattt	tggaaaaaga	tgtgataggt	ggcaagcacc	tctggcacgg	caaatacggg	11400
gtagaagttg	aggcgcggtg	tgggctcgca	tgtgccgttt	tcttggcggt	tggggggtac	11460
gcgcggtgag	aacaggtggc	gttcgtaggc	aaggctgaca	tccgctatgg	cgaggggcac	11520
atcgctgcgc	tcttgcaacg	cgtcgcagat	aatggcgcac	tggcgctgca	gatgcttcaa	11580
cagcacgtcg	tctccacat	ctaggtagtc	gccatgcctt	tggtccccc	gcccgaactg	11640
ttcctcgttt	gctcctgctg	ctccttttta	tctctgttg	tctctgttg	gtactgagcg	11700
atcctcgctg	tcttcgctta	caaaacctgg	gtcctgctcg	ataatcactt	cctcctcctc	11760
aagcgggggt	gcctcgacgg	ggaagggtgg	aggcgcggtg	gcggcatcgg	tggaggcggt	11820
ggtggcgaac	tcaaaggggg	cggttaggct	gtcctccttc	tgcactgact	ccatgatctt	11880
tttctgccta	taggagaagg	aatggccag	tcgggaagag	gagcagcgcg	aaaccacccc	11940
cgagcgcgga	cgcggtgcgg	cgcgacgtcc	accaaccatg	gaggacgtgt	cgcccccgct	12000
gccgtgcgcg	ccgcctcccc	gcgcgcccc	aaaaaagcgg	ctgaggcggc	gtctcgagtc	12060
cgaggacgaa	gaagactcgt	cacaagatgc	gctgggtgcc	cgcacacca	gccccgccc	12120
atcgacctcg	acggcggtat	tggccattgc	gtccaaaaag	aaaaagaagc	gccccctctc	12180
caagcccag	cgcccccat	ccccagaggt	cctcgtggac	agcgagggaag	aaagagaaga	12240
tgtggcgcta	caaatgggtg	gtttcagcaa	cccaccggtg	ctaatacagc	acggcaaggg	12300
aggtaagcga	acggtgcggc	ggctgaatga	agacgacca	gtggcgcggg	gtatgcggac	12360
gcaagaggaa	aagggaagag	ccagtgaagc	ggaaagtga	agcacggtga	taaaccgct	12420
gagcctgccg	atcgtgtctg	cgtgggagaa	gggcatggag	gctgcgcgcg	cgttgatgga	12480
caagtaccac	gtggataacg	atctaaaggc	aaacttcaag	ctactgcctg	accaagtgga	12540
agctctggcg	gccgtatgca	agacctggct	aaacgaggag	caccgcggtg	tgacgtgac	12600
cttcaccagc	aacaagacct	ttgtgacgat	gatggggcga	ttcctgcagg	cgtacctgca	12660

gtcgttttgca	gaggtaacct	acaagcacca	cgagcccacg	ggctgcgcgt	tgtggctgca	12720
ccgctgcgct	gagatcgaag	gcgagcttaa	gtgtctacac	gggagcatta	tgataaataa	12780
ggagcacgtg	attgaaatgg	atgtgacgag	cgaaaacggg	cagcgcgcg	tgaaggagca	12840
gtctagcaag	gccaaagatcg	tgaagaaccg	gtggggccga	aatgtggtgc	agatctccaa	12900
caccgacgca	aggtgctgcg	tgcattgacg	ggcctgtccg	gccaatcagt	tttccggcaa	12960
gtcttgcggc	atgttcttct	ctgaaggcgc	aaaggctcag	gtggctttta	agcagatcaa	13020
ggctttcatg	caggcgctgt	atcctaacgc	ccagaccggg	cacggtcacc	ttctgatgcc	13080
actacggtgc	gagtgcaact	caaagcctgg	gcatgcaccc	tttttgggaa	ggcagctacc	13140
aaagttgact	ccgttcgccc	tgagcaacgc	ggaggacctg	gacgcggatc	tgatctccga	13200
caagagcgctg	ctggccagcg	tgcaccaccc	ggcgctgata	gtgttccagt	gctgcaaccc	13260
tgtgtatcgc	aactcgcgcg	cgcagggcgg	aggccccaac	tgcgacttca	agatatcggc	13320
gcccagacgtg	ctaaacgcgt	tggatgatgt	gcgcagcctg	tggagtgaag	acttcaccga	13380
gctgcccggg	atgggttgtg	ctgagtttaa	gtggagcact	aaacaccagt	atcgcaacgt	13440
gtccctgccca	gtggcgcata	gcgatgcgcg	gcagaacccc	tttgattttt	aaacggcgca	13500
gacggcaagg	gtggggggta	aataatcacc	cgagagtgtg	caaataaaaa	catttgcctt	13560
tattgaaagt	gtctcctagt	acattatttt	tacatgtttt	tcaagtgaac	aaaagaagtg	13620
gcgctcctaa	tctgcgcact	gtggctgcgg	aagtagggcg	agtggcgctc	caggaagctg	13680
tagagctggt	cctgggttgcg	acgcaggggtg	ggctgtacct	ggggactggt	aagcatggag	13740
ttgggtaccc	cggtaataag	gttcatgggtg	gggttgtgat	ccatgggagt	ttggggccag	13800
ttggcaaaag	cgtggagaaa	catgcagcag	aatagtccac	agcgggccga	gttggggccc	13860
tgcacgcctt	gggtggactt	ttccagcggt	atacagcggt	cggggggaaga	agcaatggcg	13920
ctacggcgca	ggagtgcact	gtactcaaac	tggtaaacct	gcttgagtcg	ttggtcagaa	13980
aagccaaagg	gctcaaagag	gtagcatgtt	tttgagcgcg	ggttccaggc	aaaggccatc	14040
cagtgtacgc	ccccagtcct	gcgaccggcc	gtattgacta	tggcgccaggc	gagcttgtgt	14100
ggagaaacaa	agcctggaaa	gcgcttgtca	taggtgcccc	aaaaatatgg	cccacaacca	14160
agatctttga	caatggcctt	cagttcctgc	tcactggagc	ccatggcgcg	agctgttgtt	14220
gatgtttgct	gcttctttta	tgttgtggcg	ttgccggccg	agaagggcgt	gcgcaggtag	14280
acggctctcga	tgacgcgcgc	gtgcggctgg	tgcacacgga	ccacgtcaaa	gacttcaaac	14340
aaaacataaa	gaaggggtggg	ctcgtccatg	ggatccacct	caaaagtcac	gtctagcgcg	14400
tgggcgagg	tggcgtagag	aagggttttg	cccaggctctg	tgagtgcgcc	catggacata	14460
aagtctactgg	agaatgggat	gcgccaagg	gtcgcgacgc	aaagaaactt	tttctgggta	14520
atactgtcaa	ccgcggtttt	gcctattagt	gggtagggca	cgttggcggg	gtaagcctgt	14580
ccctcgcgca	tgggtgggagc	gaggtagcct	acgaatcctg	agttgttatg	ctggtgaaga	14640
attccaaacct	gctgatactc	cttgtattta	gtatcgctca	ccacttgccg	gctcatgggc	14700
tggaagtttc	tgaagaacga	gtacatgcgg	tccttgtagc	tttctggaat	gtagaagccc	14760
tggtagccaa	tttgttagtt	ggccaacatc	tgcaccagga	accagtccct	ggtcatgttg	14820
cactgagcta	cgttgttagcc	ctccccgtca	actgagcggt	taatctcaaa	ctcattggga	14880
gtaagcaggc	ggtcgttgcc	cggccagcta	acagaagagt	caaaggtaat	ggccaccttc	14940
ttaaagggtg	gattaagata	gaaggttccg	tcaaggatatg	gtatggagcc	agagtaggtg	15000
tagtaagggt	cgtagcctga	tcccaggga	ggggtttcct	ttgtcttcaa	gcgtgtgaag	15060
gccaacccgc	gaaatgctgc	ccagttgcgc	gatgggatgg	agatgggcac	gttgggtggcg	15120
ttggcgggta	tgggggtatag	catgttggcg	gcggaaagg	agtcattaaa	ggactggtcg	15180
ttgggtgcat	ttctgagcat	ggcttccagc	gtggaggccg	tgttgtgggc	catggggaag	15240
aagggtggcgt	aaagacaaat	gctgtcaaac	ttaatgctag	ccccgtcaac	tctaagatcg	15300
tttcccagag	agctctgcag	aaccatgtta	acatccttcc	tgaagttcca	ttcatatgta	15360
tatgagcctg	gcaggaggag	gaggttttta	atggcaaaaa	acttttgggg	cacctgaatg	15420
tgaaagggca	cgtagcggcc	gtttcccaac	aacatggagc	gataacggag	gccccgattg	15480
cgggtgggtgt	taaagggatt	aacgttgtcc	atgtagtcca	gagaccagcg	cgccccaaag	15540
ttaatgtagc	agtctacaag	cccgggagcc	accactcgct	tgttcatgta	gtcgtagggtg	15600
ttgggggtgt	cagatatattc	cacattgggtg	gggttgtatt	ttagcttgtc	tggcaggtac	15660
agcgcaatat	tggagtaaaag	gaaatttctc	catagggttg	catttaggtt	aatttccatg	15720
gcaaagggtgt	taccactcc	tatttccatta	cgtgttgcaa	aagtttcatc	ttttgtccat	15780
gtagtatctc	cattatcgcc	tgagccattg	ccattagcct	taatagcttg	ataggtgtca	15840
gttaccceaa	tacccccaaag	aggaaaacaa	taatttggca	attcatcctc	agttccatgg	15900
ttttcaatga	ttctaacatc	tggatcatag	ctgtctacag	cctgattcca	catagaaaaa	15960
tatctggttc	tatcacctat	ggaatcaagc	aagagttgat	aggacagctc	tgtgtttctg	16020
tcttgcaaat	ctaccacggc	atttagctgc	gatgcctgac	cagcaagaac	acccatgttg	16080
ccagtgtgtg	tataatacat	taggccaata	aaattgtccc	tgaagcaat	gtaattgggtg	16140

ctgtttggca	tagattgttg	acccaacata	gctttagaat	tttcatcacc	ttttccaggt	16200
ttgtaagaca	gatgtgtgtc	tggggtttcc	atattttacat	cttccactgta	caaaaccact	16260
tttggttttag	tagcattgcc	ttgccggtcg	ttcaaagagg	tagtattttga	gaagaattgc	16320
aagtcaacct	ttggaagagg	cacccctttt	tcacccggaa	ccagaacgga	ttgaccacca	16380
aaaggatttg	taggcctggc	ataagatcca	tagcatgggt	tcattgggagt	tgttttttta	16440
agcactctcc	ctcctgccgc	attagcatca	gcttcgttcc	actgagattc	gccaatttga	16500
ggttctgggt	gataggaagg	atctgcgtat	acaggtttag	cttgtgtttc	tgcatgtctc	16560
gatcctattt	gtagcccgtc	ttttgttaatt	gtttctccag	acaaaggagc	ctgggcatag	16620
acatgtgttt	tcttagtagc	ctgatctcga	gcgttttgct	cttcttcttc	ctcttcttca	16680
tcttcatctt	cctcttcttc	atcctcggca	actgcccggc	cgctatcttc	ggtttgttcc	16740
cactcacagg	agttaggagc	gcccttggga	ctagagcgt	tgtaggcagt	gccggagtag	16800
ggcttaaaag	taggccccct	gtccagcacg	ccgcggatgt	caaagtacgt	ggaagccata	16860
tcaagcacac	ggttgtcacc	cacagccagg	gtgaaccgcg	ctttgtacga	gtacgcggta	16920
tctcgcgggt	ccacagggat	gaaccgcagc	gtcaaaccgt	gggaccgggtc	tgtaggttacg	16980
tcgtgcgtag	gtgccaccgt	ggggtttcta	aacttgttat	tcaggctgaa	gtacgtctcg	17040
tggtgcgggg	caaaactgcac	cagcccgggg	ctcaggtact	ccgaggcgtc	ctggcccagag	17100
atgtgcatgt	aagaccactg	cggcatcatc	gaaggggtag	ccatcttggga	aagcgggcgc	17160
acggcgggtc	agcagctcct	ctggcggcga	catggacgca	tacatgacac	atacgacacg	17220
ttagctattt	agaagcatcg	tccgcgcttc	agggattgca	ccccagacc	cacgatgctg	17280
ttcagtgtgc	tttgcaggtt	gccactggct	acgggcccga	tcgatcgcgg	accgctggcg	17340
gcacggcgca	gggacgcgcg	ctagggcgcg	gttacacaaa	cggcggacgg	ccctggcagc	17400
acagggtttc	gctgggtgtc	agcgggggga	ggcaggtcca	gcgttacagg	tggtgtctgg	17460
cccagcactc	cggtagccat	gggcgcgatg	ggacgggtgg	tgggcaggcc	ttgcttttagt	17520
gcctcctcgt	acgagggagg	ctcatctatt	tgcgtcacca	gagtttcttc	cctgtcgggc	17580
gcgggacgct	tttcgccacg	cccctctgga	gacactgtct	ccacggccgg	tggaggctcc	17640
tctacgggag	ggcggggatc	aagcttactg	ttaatcttat	tttgactctc	ctggttggcc	17700
aggtccacca	cccgccta	gccagaggcc	aggccatcta	ccaccttttg	ttggaaattt	17760
tgctctttca	acttgtccct	cagcatctgg	cctgtgctgc	tgttccaggc	cttgctgcca	17820
tagttcttaa	tggtggaacc	gaaattttta	atgccgctcc	acagcagacc	ccagctgaag	17880
gcgccaccgc	tcataattgt	ggtgccgata	tcttgccagt	ttcccatgaa	cgggcgcgag	17940
ccgtgtcgcg	gggcccagaga	cgcaaagtgt	atgtcttcca	ttctacaaaa	tagttacagg	18000
accaagcgag	cgtgagactc	cagacttttt	attttgattt	ttccacatgc	aacttgtttt	18060
taatcagtgt	ctctgcgcct	gcaaggccac	ggatgcaatt	ccgggcacgg	cgccaatcgc	18120
cgcggcgatc	agtggaaata	ggaggggcag	gataccgccg	cgcattgcgac	ggtgcgacgc	18180
gcgccgcgcg	cgggtggtgcg	cacgacgcac	gcgcgccgtc	aggccgtggc	cggccatgcc	18240
cctcctacgg	gtcattcttc	ctcggaatcc	cggcacgggg	aaacggaggc	ggcaggtgag	18300
ggccatatct	tgaagaacca	caaagaccgg	cttttaaacg	atgctgggggt	ggtagcgcgc	18360
tgttggcagc	accagggtcc	tgctccttc	gcgagccacc	ctgcgcacgg	aaatcggggc	18420
cagcacgggc	tggcgacggc	gacggcggcg	gcgggttcca	gtggtggttc	ggcgtcgggt	18480
agtcgctcgt	cttctggggc	ggtaggtgta	gccacgatag	ccgggggtag	gcgcgatgga	18540
aggatgtagg	gcataatcgg	gcagtagtgc	ctggcggtg	ccgtacttcc	tggaaacggc	18600
cgggcgcggc	ggggctgaaa	cgcgaaacat	ccacgggtcc	gtttgcacct	ccgtagaggt	18660
tttggaacgcg	gccgcagcgg	cgcctgcac	cgcggcatct	gccaccgccg	aggcaaccgg	18720
ggacgtttgt	gtctccatgc	cctctgtggc	agtggcaata	ctagtgtctac	tggtggtggg	18780
tatctgaacg	tccacggtct	gcacgccag	tccgggtgcc	acctgcttga	ttggccgcac	18840
gcggacctcg	ggctccagcc	caggctccac	ggtcattttt	tccaagacat	cttcagtcg	18900
ctggcgcttg	ggtaccatca	ctgcacgggt	gggtgccaag	tcaccagact	cgcgctttag	18960
gccgcgcttt	tcttcggacg	gtgcaagcgt	gggcagcacc	tgctgcagtg	tcacgggctt	19020
taggctagggt	ggtgggttgc	cctcgtccag	cggcaacgcc	aacatgtcct	tatgccgctt	19080
tccgtaggga	aactccccga	ggcgtcgtt	ggcctgtca	agcaggtcct	cgctgcgcta	19140
cacctcatca	tacacgcgct	tgtaggtgcg	ggtggagcgc	tcaccgggcg	taaaaactac	19200
ggtggtgccc	ggtcgcaaaa	cacgtcttac	gcgtcgacct	ttccactgta	cccgcgcct	19260
gggcgcgggt	gcgtgcagca	gttccacctc	gtcgtcaagt	tcacatcat	catcatcttt	19320
ctttttcttt	ttgacctgct	ttagctttcg	gggcttgtaa	tcctgctctt	ccttcttcgg	19380
ggggccatag	atctccggcg	cgatgacctg	gagcatctct	tctttgattt	tgcgcttgga	19440
catagcttgc	ttgcgcggcg	cgcgcgctgg	atatatacaa	cagtacgagt	ctaagtagtt	19500
ttttcttgca	atctagttgc	cggggggcg	gttcgcacgc	ggcacgcgca	ggcgcgtaac	19560
cgagtgcgcg	acccagtaca	cgttgccctc	gcgacctga	gtcatagcac	taatggccgc	19620

```

ggctgctgcg ggggcccgtc gtcgcctgga cctggggggc acagtgacaa taccgcggc 19680
cagccttcga gggcccgca tggccgcccg tggccgggtg cgacgtgcgc ggtaagcag 19740
ggccgcccgc gcgcgttggg cggcagtgcc gggtcggcgg cgggtggcgac gtgctacgcg 19800
cctccgcccgt ctcttcattt tagcataacg cggggtcccg cgcaccacgg tctgaatggc 19860
cgcgtccact gtggacactg gtggcgccgt gggcggttag ttgcgcgcct cctccaccac 19920
cgctcaatg gcgtcatcga ggggtggtcg ccagtgcgcg ccgcgtttgt gcgcgcccc 19980
gggcgcgcggt tagtgcccgc gcacgcgcac tgggtgttgg tggagcgct tctttgcccc 20040
gccaacatc ttgcttgga agcgcaggcc ccagcctgtg ttattgctgg gcgatataag 20100
gatggacatg tttgctcaaa aagtgcggct cgataggacg cgcggcgaga ctatgccag 20160
ggccttghaa acgtaggggc aggtgcggcg tctggcgta gtaatggtca ctgcgtggac 20220
tcctccgatg ctggttcgca cgggtagcgt cccgtgatct gtgagagcag gaacgttttc 20280
actgacgggtg gtgatggtgg gggctggcgg gcgcgcaaaa atctggttct cgggaaagcg 20340
attgaacacg tgggtcagag aggtaaactg gcggtagagc tgggagtaga cggcctgggtc 20400
gtttagaag ctcttgaggt gcacgggcaa cagctcggcg cccaccaccg gaaagttgct 20460
gatctggctc gtggagcgga aggtcacggg gtcttgcatc atgtctggca acgaccagta 20520
gacctgctcc gagccgagg ttacgtcagg agtgcaaaagg aggtccatg agcgcatccc 20580
ggcttgaggg tcgccgtagt tgtatgcaag gtaccagctg cggtagtggg tgaaggtgct 20640
gtcattgctt attaggttgt aactgcgttt cttgctgtcc tctgtcaggg gtttgatcac 20700
cggtttcttc tgaggcttct cgacctcggg ttgcgcagcg ggggcggcag cttctgccgc 20760
tgcttcggcc tcagcgcgct tctctccgc cgtgttggca aaggtgtcgc cggaatggc 20820
atgatcgttc atgtcctcca ccgctgcac tgcgcggct gccgcgttg agttctcttc 20880
cgcccgctg ccactgttgt tgcgcgcgc tgcgccatcc ccgcctgtt cgggtgtcatc 20940
ttttaagctt gcctggtagg cgtccacatc caacagtgcg ggaatgttac caccctccag 21000
gtcatcgtag gtgatcctaa agcctcctg gaagggttgc cgcttgcgga tgcccaacaa 21060
gttgctcagg cggctgtggg tgaagtcac cccgcaccc ggcagcaaaa tgatgtctg 21120
atggaaggct tcgtttgtat ataccacagg catgacaaga ccagtgaact ggtaaaccc 21180
cagctgaag ttgcgggtgt caaactttac ccgatgtcg ctttcagaa ccccgttctg 21240
cctgcccact ttcaagtagt gctccacgat cgcgttgttc ataaggtcta tggatcatg 21300
ctcggagtag ttgccctcgg gcagcgtgaa ctcacccac tcatatttca gctccacctg 21360
tttgcctta gtaagcgagc ggcacaccat caccgcgc ttaaacttat tggtaaacat 21420
gaactcgctt acatttgga ttttggtatg caggatggtt ttcaggtcgc cgcccagtg 21480
cgaacggtcg tcaagattga tggctgtgt gcttgccctc cccgggctgt agtcattgtt 21540
ttgaatgacc gtggttagaa agttgctgtg gtcgttctgg tagttcagg atgccacatc 21600
cgttgacttg ttgtccacaa ggtacacacg ggtggtgtcg aataggggtg ccaactcaga 21660
gtaacggatg ctgtttctcc ccccggtagg ccgcaggta cgcggaggca caaacggcg 21720
gtcagggga gcatcgaagg gggaaccag cgcgcgcgc actggcgcc cgctcaccac 21780
gctctcgtag gagggaggag gacctctc atacatcgcc gcgcgtgca tactaagggg 21840
aatacaagaa aaccaacgct cgggtccatg gccttggtga gttttttatt ttgcatcatg 21900
ctttttttt ttttttttaa aacattctcc ccagcctggg gcgaagggtg gcaaacgggt 21960
tgccactccc tcccaaatcc aggcagctgc tgtcgtctgc cgagtcacg tcctcccaca 22020
ccagaccccg ctgacggtcg tgcctttgac cggcggtgg cggcgcgcg ccgggacat 22080
ccctgtgctc ctgcgcatac gtcttccatc tactcatctt gtccactagg ctctctatcc 22140
cgttgttggg aatgcccga ggcaggttct tttcgcgctg cggctgcagc agcgagttgt 22200
ttaggtactc ctccctcgccc agcaggcgcg ggcgggtggt gcgagtgctg gtaaaagacc 22260
ctatcaagct tggaaatggg ctactcgcat ctgaccgcgg ggcgcagcg cctagatcgg 22320
acaagctgct tggcctgcgg aagctttcct ttcgcagcg cgcctctgcc tgcctcgct 22380
gttgcaactc tagcagggtc tgcggttcg gggaaaacac gctgtcgtct atgtcgtccc 22440
agaggaatcc atcgttacc tgggcacct caaatcccc ggtgtagaaa ccagggggcg 22500
gtagccagtg cgggttcaag atggcattgg tgaaatactc ggggttcacg gcggccgcgc 22560
gatgcaagta gtccattagg cgattgataa acggcgggtt tgaggcatac atgcccgggt 22620
ccatgttgcg cgcggtcatg tccagcgcca cgttgggcgt taccgcgtc cgcacagg 22680
taaggctcac gctctgctgc acatagcgca agatgcgctc ctcctcgctg tttaaactgt 22740
gcaacgaggg gatcttctgc cgcgggttg tcagcaggta gttcagggtt gcctccaggc 22800
tgcccgtgtc ctccctgccc agcgcgcggc tgacacttgt aatctcctgg aaagtatgct 22860
cgtccacatg cgctgacct atggcctcgc ggtacagtgt cagcaagtga cctaggtatg 22920
tgtccggga cagctgcca ctgtccgtga agggcgctat tagcagcagc aacaggcgcg 22980
agttggcggt cagcaagcta gacaggtcg cgggtcgcc tgtgggagcc cgcaccccc 23040
acagcccctg caagttcttg aaagcctggc tcaggtttac ggtctgcagg cttgtctac 23100

```

tggctctggaa	aaaatagtct	ggccccgact	ggtacacctc	actttgcggt	gtctcagtca	23160
ccattagccg	cagtgcgctc	acaaagttag	tgtagtccct	ctgtccccgc	ggcacgttgg	23220
cgggctgtgt	actcaggaag	gcgttttagt	caaccatgga	gcccaggttg	ccctgctgct	23280
gcgcgcgctc	acgctgcgcc	acggcctcgc	gcacatcccc	caccagccgg	tccaggttgg	23340
tctgcacggt	gccgctgttg	taacgagcca	cgcgctgaag	cagcgcgtcg	tagaccaggc	23400
cggcctcctc	gggccggatg	gccctgtttt	cggccagcgc	gtttacgata	gccagcacct	23460
tctcgtgcgt	ggggtttgcg	cgcgcgggga	ccaccgcttc	cagaattgcg	gagagccggt	23520
tggcctgcgg	ctgctgccgg	aacgcgtcag	ggttacgcgc	agtcagcgac	atgatgcggt	23580
ccatgacctg	gcgccagtcg	tccgtggagt	taaggccgga	cggctggctc	tgcagcgccg	23640
cccgcaccgc	cgggtccggt	gcgtcttgca	tcattctgac	agaaacatca	ccgcttagta	23700
ctcgcgctcc	tctggctcgt	actcatcgct	ctcgtcatat	tcctccacgc	cgcgcagctt	23760
gccagcgcgc	gcgggtgcca	cgcgcagccc	aggtccggcc	ccagctgcct	ccagggcgcg	23820
tccgcttggg	gcccagcgca	ggtcagcgcc	cgcgtcaaa	taggactcgg	cctctctatc	23880
gccgctgccc	gtgccagcca	gggccctttg	caggctgtgc	atcagctcgc	ggtcgtgag	23940
ctcgcgcgcg	cggctcacgc	tcacggcctt	gtggatgcgc	tcgttgcgat	aaacgcccag	24000
gtcgtcgctc	aaggtaaagca	ccttcaacgc	catgcgcata	tagaaccctt	cgatctttac	24060
ctccttgtct	atgggaacgt	aaggggtatg	gtatatcttg	cgggcgtaaa	acttgcccag	24120
actgagcatg	gaatagttaa	tggcggccac	cttgtcagcc	aggctcaagc	tgcgctcctg	24180
caccactatg	ctctgcagaa	tgtttatcaa	atcgagcagc	cagcggccct	cgggctctac	24240
tatgttttagc	agcgcattccc	tgaatgcctc	gttgtccctg	ctgtgctgca	ctataaggaa	24300
cagctgcgcc	atgagcggct	tgtctatttg	gttgtctccc	agcgcgctta	caaagtccca	24360
cagatgcata	agtcctatag	ccacctcctc	gcgcgccaca	agcgtgcgca	cgtgggttgt	24420
aaagcttttt	tgaagtttaa	tctcctggtt	caccgtctgc	tcgtacgcgg	ttaccaggtc	24480
ggcggccgcc	acgtgtgcgc	gcgcgggact	aatcccggtc	cgcgcgtcgg	gctcaaagtc	24540
ctcctcgccg	agcaaccgct	cgcggttcag	gccatgcgcg	aactcgcgcg	ctgcgtggaa	24600
ctttcgatcc	cgcattctct	cgggctcctc	tccctcgcgg	tcgcgaaaaca	ggttctgccc	24660
cggcacgtac	gcctcgcgcg	tgtcacgctt	cagctgcacc	cttgggtgtc	gctcaggaga	24720
ggcgctcct	agccgcgcca	ggccctcgcc	ctcctccaa	tccaggtagt	gccgggcccc	24780
gcgcgcgggg	ggttcgtaat	caccatctgc	cgcgcgctca	gccgcggatg	ttgccccctc	24840
tgacgcggta	ggagaagggg	aggggtgccct	gcattgtctg	cgtgtctctt	gctcttgccg	24900
ctgctgagga	ggggggcgca	tctgccgcag	caccggatgc	atctgggaaa	agcaaaaaag	24960
gggctcgtcc	ctgtttccgg	aggaatttgc	aagcggggtc	ttgcatgacg	gggaggcaaa	25020
cccccgcttc	ccgcagtcgg	gccggcccca	gactcgaacc	gggggtcctg	cgactcaacc	25080
cttgaaaaat	aacctccggg	ctacagggag	cgagccactt	aatgctttcg	ctttccagcc	25140
taaccgctta	cgcgcgcgcg	ggccagtggc	caaaaaagct	agcgcagcag	ccgcgcgcgc	25200
tggaaaggag	ccaaaaggag	cgtcccccgc	ttgtctgacg	tcgcacacct	gggttcgaca	25260
cgcggggcgg	aaccgcatgg	atcacggcgg	acggccggat	ccgggggttc	aaccccggtc	25320
gtccgccatg	atacccttgc	gaatttatcc	accagaccac	ggaagagtgc	ccgcttacag	25380
gctctccttt	tgcacggtct	agagcgtcaa	cgactgcgca	cgcctcaccg	gccagagcgt	25440
cccgaccatg	gagcactttt	tgcgcgtcgc	caacatctgg	aaccgcgtcc	gcgactttcc	25500
gcgcgcctcc	accaccgccc	cgggcatacc	ctggatgtcc	aggtaacatc	acggatatca	25560
tcgccttatg	ttggaagacc	tcgcccccg	agccccggcc	accctacgct	ggccccctca	25620
ccgccagccg	ccgccgcact	ttttggtggg	atatcagtac	ctgggtgcgga	cttgcaacga	25680
ctacgtcttt	gactcaaggg	cttactcgcg	tctcaggtac	accgagctct	cgcagccggg	25740
tcaccagacc	gttaactggt	ccgttatggc	caactgcact	tacaccatca	acacgggcgc	25800
ataccaccgc	tttgtggaca	tggatgactt	ccagtctacc	ctcacgcagg	tgcagcaggc	25860
catattagcc	gagcgcgttg	tcgccgacct	ggccctgctt	cagccgatga	ggggcttcgg	25920
ggtcacacgc	atgggaggaa	gagggcgcca	cctacggcca	aactccgccc	ccgccgtagc	25980
gatagatgca	agagatgcag	gacaagagga	aggagaagaa	gaagtgcggg	tagaaaggct	26040
catgcaagac	tactacaaa	acctgcgccg	atgtcaaaac	gaagcctggg	gcattggccga	26100
ccgctgcgcc	attcagcagg	cgggacccaa	ggacatgggt	cttctgtcga	ccatcccgcc	26160
tctcaagacc	gcctacttta	attacatcat	cagcagcacc	tcgcgcagaa	acaacccccg	26220
ccgccacccg	ctgcgcgccg	ccacggtgct	cagcctacct	tgcgactgtg	actggttaga	26280
cgcctttctc	gagaggtttt	ccgatccggg	cgatgcggac	tcgctcaggt	ccctcggtgg	26340
cggagtacct	acacaacaat	tgttgagatg	catcgtttag	gccgtatccc	tgcgcgacgg	26400
cagccccccg	ccaaccata	accgggacat	gacgggcggc	gtcttccaac	tgcgcccccg	26460
cgaagaacgc	cgcgcggtca	cgcgcgctcg	gcgcgcgctc	gcgggggaga	tgatcgagcg	26520
ctttgtcgac	cgcctcccg	tgcgcgctcg	tcgcgcgctc	gtccccctc	ccccaccgcc	26580

gccagaagaa	gaagaagaag	gggaggccct	tatggaagag	gagattgaag	aagaagaggc	26640
ccctgtagcc	tttgagcgcg	aggtgcgcg	cactgtcgcc	gagctcatcc	gtcttctgga	26700
ggaggagtta	accgtgtcgg	cgcgcaactc	ccagtttttc	aacttcgccc	tggacttcta	26760
cgaggccatg	gagcgccctg	aggccttggg	ggatatcaac	gaatccacgt	tgcgacgctg	26820
ggttatgtac	ttcttcgtgg	cagaacacac	cgccaccacc	ctcaactacc	tctttcagcg	26880
cctgcgaaac	tacgccgtct	tcgcccggca	cgtggagctc	aatctcgcg	aggtgggtcat	26940
gcgcgcccgc	gatgccgaag	ggggcggtgt	ctacagccgc	gtctggaacg	agggaggcct	27000
caacgccttc	tcgcagctca	tggcccgcct	ctccaacgac	ctcgccgcca	ccgtggagcg	27060
agccggagcg	ggagatctcc	aggaggaaga	gatcgagcag	ttcatggccg	aaatcgcccta	27120
tcaagacaac	tcaggagacg	tgcaggagat	tttgcgccag	gcccgcgtca	acgacaccga	27180
aattgattct	gtcgaactct	ctttcagggt	caagctcacc	gggcccgtcg	tcttcacgca	27240
gaggcgccag	attcaggaga	tcaaccgccc	cgtcgtcgcg	ttcgccagca	acctccgccc	27300
gcagcaccag	ctcctgcccg	cgcgcgccgc	cgacgtgccc	ctgccccctc	tcccggcggg	27360
tcccagagccc	cccctacctc	cggggggccc	cccgcgtcac	cgtttttaga	tgcacatccc	27420
aaggacaccc	ccgcgcccca	ccgcccgcgc	cgcggtaccg	tagtcgcgcc	gcggggagct	27480
ggctctttgc	aagtcacatc	cgccgcccac	aaccagcccc	tggaaatcag	gtatcacctg	27540
gacctagccc	gcgccttgac	ccggctatgc	gaggtaaacc	tgcaggagct	cccgcctgac	27600
ctgtcgccgc	gggagctcca	gacctgggac	agctcccac	tgcgcgatgt	tgtcatcaag	27660
ctccgacccg	cgcgcgccga	catctggact	ttgggctcgc	gcggcggtgt	ggtccgatcc	27720
accataaact	ccctcgagca	gccagacggt	caaggacaag	cagccgaagt	agaagaccac	27780
cagccaaacc	cgccaggcga	ggggctcaaa	ttcccactct	gcttccttgt	gcgcggtcgt	27840
cagggtcaacc	tcgtgcagga	tgtacagccc	gtgcaccgct	gccagtactg	cgcacgtttt	27900
tacaaaagcc	agcacgagtg	ttcgcccggt	cgccaggact	tctactttca	ccacatcaac	27960
agccactcct	ccaactgggtg	gcgggagatc	cagttcttcc	cgatcggtc	gcacctcgc	28020
accgagcgct	tctttgtcac	ctacgatgta	gagacctata	cttggtggg	ggcctttggg	28080
aagcagctcg	tgcccttcac	gctggttatg	aagttcgccg	gagatgagcc	tctggtgacc	28140
gccgcgagag	acctagccgt	ggaccttgga	tgggaccgct	gggaacaaga	cccgcttacc	28200
ttctactgca	tcacccacga	aaaaatggcc	ataggtcgcc	agtttaggac	ctttcgcgac	28260
cacctgcaaa	tgctaattgg	ccgtgacctg	tggagctcat	tcgtcgcttc	caaccctcat	28320
cttgacagct	gggccttgct	agaacacggg	ctcagctccc	ctgaggagct	caacctacag	28380
gaacttaaaa	aattgccttc	catcaagggc	accccgcgct	tcttggaact	ttacatcggt	28440
ggccacaaca	tcaacggctt	cgacgagatc	gtgctcgccg	cccaggtaat	taacaaccgt	28500
tccgagggtc	cgggaccctt	ccgcatcaca	cgcaacttta	tgccctcgcg	gggaaagata	28560
cttttcaacg	atgtcacctt	cgccctgcca	aaccgcggtt	ccaaaaagcg	cacggacttt	28620
ttgctctggg	agcagggcgg	atgcgacgac	actgacttca	aataccagta	cctcaaagtc	28680
atggttaggg	acacctttgc	gctcaccac	acctcgctcc	ggaaggccgc	gcaggcatac	28740
gcgtaccgg	tagaaaaggg	atgctgcgcc	taccaggccg	tcaaccagtt	ctacatgcta	28800
ggctcttacc	gttcggaggc	cgacgggttt	ccgatccaag	agtactggaa	agaccgcgaa	28860
gagtttgtcc	tcaaccgcga	gctgtggaaa	aaaaagggac	aggataagta	tgacatcatc	28920
aaggaaaccc	tggactactg	cgccctagac	gtgcaggtea	ccgccgagct	ggtcaacaag	28980
ctgcgcgact	cctacgcctc	cttcgtgcgt	gacgcggtag	gtctcacaga	cgccagcttc	29040
aacgtcttcc	agcgtccaac	catatcatcc	aactcacatg	ccatcttcag	gcagatagtc	29100
ttccgagcag	agcagcccgc	ccgtagcaac	ctcggtcccg	acctcctcgc	tccctcgcac	29160
gaactatacg	attacgtgcg	cgccagcatc	cgcggtggaa	gatgctaccc	tacatatctt	29220
ggaatactca	gagagcccct	ctacgtttac	gacatttgcg	gcattgtacgc	ctccgcgctc	29280
accacaccca	tgccatgggg	tccccactc	aaccatacgc	agcgcgcgct	tgcgcgccgc	29340
gcatggcagc	agcgctaga	cttgcaagga	tgcaagatag	actacttcga	cgcgcgctcg	29400
ctgcccgggg	tctttaccgt	ggacgcagac	cccccgagc	agacgcagct	agaccacta	29460
ccgccattct	gttcgcgcaa	ggcgcgccgc	ctctgctgga	ccaacgagcg	cctacgcgga	29520
gaggtagcca	ccagcgttga	ccttgtcacc	ctgcacaacc	gcgggtggcg	cgtgcacctg	29580
gtgcccagcg	agcgaccac	cgtctttccc	gaatggcggt	gcgttgccgc	cgaatacgtg	29640
cagctaaaca	tcgcggccaa	ggagcgcgcc	gatcgcgaca	aaaaccaaac	cctgcgctcc	29700
atcgccaagt	tgctgtccaa	cgccctctac	gggtcgtttg	ccaccaagct	tgacaacaaa	29760
aagattgtct	tttctgacca	gatggacgcg	gccaccctca	aaggcatcac	cgcgggccag	29820
gtgaatatca	aatcctcctc	gtttttggaa	actgacaatc	ttagcgcgag	agtcatgccc	29880
gcttttgaga	gggagtactc	accccaacag	ctggccctcg	cagacagcga	tgcggaagag	29940
agtgaggacg	aacgcgcccc	cacccctttt	tatagcccc	cttcaggaa	accgggtcac	30000
gtggcctaca	cctataaacc	aatcaccttc	cttgatgccg	aagaggcgca	catgtgtctt	30060

cacacccctgg	agcgagtggg	ccccctagtg	gacaacgacc	gctacccctc	ccacttagcc	30120
tccttcgtgc	tggcctggac	gcgagccttc	gtctcagagt	ggtccgagtt	tctatacgag	30180
gaggaccgcg	gaacaccgct	cgaggacagg	cctctcaagt	ctgtatacgg	ggacacggac	30240
agccttttcg	tcaccgagcg	tggacaccgg	ctcatggaaa	ccagaggtaa	gaaacgcata	30300
aaaaagcatg	ggggaaacct	ggtttttgac	cccgaacggc	cagagctcac	ctggctcgtg	30360
gaatgcgaga	ccgtctgcgg	ggcctgcggc	gcggatgcct	actccccgga	atcggtatct	30420
ctcgcgcccc	agctctacgc	ccttaaaagt	ctgcactgcc	cctcgtgcgg	cgcctcctcc	30480
aagggcaagc	tgcgcgcccc	gggccacgcc	gcggaggggc	tggactatga	caccatggtc	30540
aaatgctacc	tggccgacgc	gcaggggcaa	gaccggcagc	gcttcagcac	cagcaggacc	30600
agcctcaagc	gcaccctggc	cagcgcgcag	cccggagcgc	accccttcac	cgtgaccagc	30660
actacgctga	cgaggaccct	gcgcccgtgg	aaagacatga	ccctggcccc	tctggacgag	30720
caccgactac	tggcgtacag	cgaaagccgc	cccaaccgcg	gaaacgagga	gatatgctgg	30780
atcgagatgc	cgtagagcac	gtgaccgagc	tgtgggaccg	cctggaaactg	cttgggtcaaa	30840
cgctcaaaaag	catgcctacg	gcggacggcc	tcaaaccggt	gaaaaacttt	gcttccttgc	30900
aagaactgct	atcgctgggc	ggcgagcgcc	ttctggcgca	tttggtcagg	gaaaacatgc	30960
aagtcaggga	catgcttaac	gaagtggccc	ccctgctcag	ggatgacggc	agctgcagct	31020
ctcttaacta	ccagttgcag	ccggtaatag	gtgtgatctt	cgggcccacc	ggctgcggta	31080
agtcgcagct	gctcaggaac	ctgctttctt	cccagctgat	ctccccctacc	ccggaacagg	31140
ttttcttcat	cgccccgcag	gtagacatga	tccccccatc	tgaactcaaa	gcgtgggaaa	31200
tgcacatctg	tgagggtaac	tacgccccctg	ggccggatgg	aaccattata	ccgcagctctg	31260
gcacccctcg	cccgcgcttt	gtaaaaatgg	cctatgacga	tctcatcctg	gaacacaaatg	31320
atgacgttag	tgatcccaga	aatatcttcg	cccaggccgc	cgccccgtggg	cccattgcca	31380
tcattatgga	cgaatgcatg	gaaaatctcg	gaggtcacia	gggcgtctcc	aagttcttcc	31440
acgcattttcc	ttctaagcta	catgacaaat	ttcccaagtg	caccggatac	actgtgctgg	31500
tggttctgca	caacatgaat	ccccggaggg	atatggctgg	gaacatagcc	aacctaaaaa	31560
tacagtccaa	gatgcattct	atatccccac	gtatgcaccc	atcccagctt	aaccgctttg	31620
taaacactta	caccaagggc	ctgccccctg	caatcagctt	gctactgaaa	gacattttta	31680
ggcaccacgc	ccagcgctcc	tgctacgact	ggatcatcta	caacaccacc	ccgcagcatg	31740
aagctctgca	gtgggtgctac	ctccaccccc	gagacgggct	tatgcccata	tatctgaaca	31800
tccagagtca	cctttaccac	gtcctggaaa	aaatacacag	gacctcaac	gaccgagacc	31860
gctggctccc	ggcctaccgc	gcgcgcaaaa	cccctaaata	aagacagcaa	gacacttgct	31920
tgatcaaaaat	ccaaacagag	tctggttttt	atttatgttt	taaaccgcat	tgggagggga	31980
ggaagccttc	agggcagaaa	cctgctggcg	cagatccaac	agctgctgag	aaacgacatt	32040
aagttcccgg	gtcaaagaat	ccaattgtgc	caaaagagcc	gtcaacttgt	catcgcgggc	32100
ggatgaacgg	gaagctgcac	tgcttgcaag	cgggctcagg	aaagcaaagt	cagtcacaa	32160
cccgcggggc	tgggctgcag	cggctgaagc	ggcggcggag	gctgcagtct	ccaacggcgt	32220
tccagacacg	gtctcgtagg	tcaaggtagt	agagtttgcg	ggcaggacgg	ggcgaccatc	32280
aatgctggag	cccacacat	tctgacgcac	cccggcccat	gggggcatgc	gcgttgtaaa	32340
atatgagctc	acaatgcttc	catcaaacga	gttgccgctc	atggcggcgg	ctgctgcaaa	32400
acagatacaa	aactacatga	gacccccacc	ttatatattc	tttcccacc	ttaagccccg	32460
cccacgatg	gcaaacagct	attatgggta	ttatgggtgc	tagcgacatg	aggttgcccc	32520
gtattcagtg	tcgctgattt	gtattgtctg	aagttgtttt	tacgttaagt	tgatgcagat	32580
caattaatac	gatacctgcg	tcataattga	ttatttgacg	tggtttgatg	gcctccacgc	32640
acgttggtgat	atgtagatga	taatcattat	cactttacgg	gtcctttccg	gtgatccgac	32700
aggttacggg	gcggcgacct	cgcgggtttt	cgtattttat	gaaaaatttc	cggtttaagg	32760
cgtttcggtt	cttcttcgtc	ataacttaat	gtttttatct	aaaataccct	ctgaaaagaa	32820
aggaaacgac	aggtgctgaa	agcgaggctt	tttggcctct	gtcgtttcct	ttctctgttt	32880
ttgtccgtgg	aatgaacaat	ggaagttaac	ggatccaggc	cgcgagcaaa	aggccagcaa	32940
aaggccagga	accgtaaaaa	ggccgcgttg	ctggcggttt	tccataggct	ccgccccctc	33000
gacgagcatc	acaaaaatca	acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	33060
agataccagg	cgtttccccc	tggaaagctc	ctcgtgcgct	ctcctgttcc	gacctgccc	33120
cttacccgat	acctgtccgc	ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	33180
cgctgtaggt	atctcagttc	ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	33240
ccccccgttc	agcccagacc	ctgcgcctta	tccggttaact	atcgtcttga	gtccaaccgc	33300
gtaagacacg	acttatcgcc	actggcagca	gccactggta	acaggattag	cagagcgagg	33360
tatgtaggcg	gtgctacaga	gttcttgaag	tgggtgccta	actacggcta	cactgaaga	33420
acagtatttg	gtatctgcgc	cttgccaaag	ccagttacct	tcggaaaaag	agttggtagc	33480
tcttgatccg	gcaaacaaa	caccgctggt	agcgggtggt	tttttggttg	caagcagcag	33540

```

attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 33600
gctcagtggg acgaaaaactc acgttaaggg attttgggtca tcagattatc aaaaaggatc 33660
ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 33720
taaacttggg ctgacagtta ccaatgctta atcagtggag cacctatctc agcgatctgt 33780
ctatttcgtt catccatagt tgcctgactc cccgtagtgt agataactac gatacgggag 33840
ggcttaccat ccggccccag tgctgcaatg ataccgctg acccacgctc accggctcct 33900
gatttatcag caataaacca gccagccgga agtgccgagc gcagaagtgg tcctgcaact 33960
ttatccgcct ccatccagtc tattagtgtg tgccgggaag cttagagtaag tagttcgcca 34020
gttaatatgt ttcgcaacgt tgttgccatt gctacaggca tcgtgggtgtc acgctcgctc 34080
tttgggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 34140
atgttggtgca aaaaagcggg tagctccttc ggtcctccga tagttgtcag aagtaagttg 34200
gccgcagtgt tatcactcat gggttatggca gcaactgcata attctcttac tgtcatgcca 34260
tccgtaagat gcttttctgt gactggtgag tattcaacca agaatacggg ataataccgc 34320
gccacatagc agaactttaa aagtgtcat cattgggaaa cgttcttcgg ggcgaaaact 34380
ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgcg caccgaagt 34440
atcttctgca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa 34500
tgccgcacaaa aaggggaataa gggcgacacg gaaatgttga atactcatac ttttctttt 34560
tcaatattat tgaagcattt atcagggtta ttgtctcatc agcgataca tatttg 34616

```

<210> 3

<211> 31672

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 3

```

gaatcgccca ggcgaatta actataacgg tcctaaggta gcgtcatcat cataatatac 60
cttattttgg attgaagcca atatgataat gagggggtgg agtttgtgac gtggcgcggg 120
gcgtgggaac gggcggggtg acgtagggtt tagggcggag taacttgcac gtattgggaa 180
ttgtagtttt tttaaaatgg gaagttacgt atcgtgggaa aacggaagtg aagatttgag 240
gaagttgtgg gttttttggc tttcgtttct gggcgtaggt tcgcgtgcgg ttttctgggt 300
gttttttgtg gactttaacc gttacgtcat ttttagtcc tataataact cgctctgtac 360
ttggcccttt ttacactgtg actgattgag ctgggtgccg gtcgagtggg gttttttaat 420
aggttttttt actggttaagg ctgactgtta tggctgccgc tgtggaagcg ctgtatgttg 480
ttctggagcg ggagggtgct attttgccta ggcaggaggg ttttccaggt gtttatgtgt 540
ttttctctcc tattaatttt gttatactc ctatgggggc tghtaatgtt tctctacgcc 600
tgcggtatag tattccccg ggtatttcg gtcgctttt agcactgacc gatgttaacc 660
aacctgatgt gtttaccgag tcttacatta tgactccgga catgaccgag gaactgtcgg 720
tggtgctttt taatcacggg gaccagttt tttacggtca cgccggcatg gccgtagtcc 780
gtcttatgct tataagggtt gtttttcctg ttgtaagaca ggcttcta atgtttaaatgt 840
ttttttttgt tattttattt tgtgtttaat gcaggaaccc gcagacatgt ttgagagaaa 900
aatggtgtct ttttctgtgg tggttccgga acttacctgc ctttatctgc atgagcatga 960
ctacgatgtg cttgcttttt tgcgcgaggg tttgcctgat tttttgagca gcaccttgca 1020
ttttatatcg ccgcccatgc aacaagctta cataggggct acgctgggta gcatagctcc 1080
gagtatgcgt gtcataatca gtgtgggttc ttttgtcatg gttcctggcg ggggaagtggc 1140
cgcgtgggtc cgtgcagacc tgcacgatta tgttcagctg gccctgcgaa gggacctacg 1200
ggatcgcggt atttttgtta atgttccgct tttgaatctt atacaggtct gtgaggaacc 1260
tgaatttttg caatcatgat tcgctgcttg aggctgaagg tggagggcgc tctggagcag 1320
atttttacaa tggccggact taatattcgg gatttgctta gagacataat gataaggttg 1380
cgagatgaaa attatttggg catggttgaa ggtgctggaa tgtttataga ggagattcac 1440
cctgaagggt ttagccttta cgtccacttg gacgtgaggg cagtttgcct tttggaagcc 1500
attgtgcaac atcttacaaa tgccattatc tgttcttttg ctgtagagtt tgaccacgcc 1560
accggagggg agcgcgttca cttaatagat cttcattttg aggttttggg taactctttg 1620
gaataaaaaa aaaaaaaca tggttcttcc agctcttccc gctcctccc tgtgtgactc 1680

```


gcagaacgaa	tgtgtagggt	ggctgggtgt	ggcttattct	gcgggtggtg	atgttatcag	1740
ggcagcggcg	catgaaggag	tttacataga	acccgaagcc	agggggcgcc	tggatgcttt	1800
gagagagtgg	atatactaca	actactacac	agagcgagct	aagcgacgag	accggagacg	1860
cagatctggt	tgtcacgccc	gcacctgggt	ttgcttcagg	aaatatgact	acgtccggcg	1920
ttccatttgg	catgacacta	cgaccaacac	gtctcgggtt	gtctcggcgc	actccgtaca	1980
gtagggatcg	cctacctcct	tttgagacag	agaccgcgc	taccatactg	gaggatcatc	2040
cgctgctgcc	cgaatgtaac	actttgacaa	tgcacaacgt	gagttacgtg	cgaggctctc	2100
cctgcagtgt	gggattttacg	ctgattcagg	aatgggttgt	tccctgggat	atgggtctga	2160
cgcgggaggga	gcttgaatac	ctgaggaagt	gtatgcacgt	gtgcctgtgt	tgtgccaca	2220
ttgatatacat	gacgagcatg	atgatccatg	gttacgagtc	ctgggctctc	cactgtcatt	2280
gttccagttcc	cggttccctg	cagtgcatag	ccggcgggca	ggttttggcc	agctggttta	2340
ggatgggtgg	ggatggcgcc	atgtttaatc	agaggtttat	atggtaccgg	gaggtggtga	2400
attacaacat	gccaaaagag	gtaatgttta	tgtccagcgt	gtttatgagg	ggtcgccact	2460
taatctacct	gcgcttgtgg	tatgatggcc	acgtgggttc	tgtgggtccc	gccatgagct	2520
ttggatacag	cgcttgcac	tgtgggattt	tgaacaatat	tgtggtgctg	tgtgcagct	2580
actgtgctga	tttaagttag	atcagggtgc	gctgctgtgc	ccggaggaca	aggcgtctca	2640
tgtcgcgggc	ggtgcgaatc	atcgctgagg	agaccactgc	catgttgtat	tcctgcagga	2700
cggagcggcg	gcggcagcag	tttattcgcg	cgctgctgca	gcaccaccgc	cctatcctga	2760
tgcacgatta	tgactctacc	cccatgtagg	cgtggacttc	cccttcgccg	cccgttgagc	2820
aaccgcaagt	tggacagcag	cctgtggctc	agcagctyga	cagcgacatg	aacttaagcg	2880
agctgcccg	ggagtttatt	aatatcactg	atgagcgttt	ggctcgacag	gaaaccgtgt	2940
ggaatataac	acctaagaat	atgtctgtta	cccatgatat	gatgcttttt	aaggccagcc	3000
ggggagaaag	gactgtgtac	tctgtgtgtt	gggaggagg	tggcaggttg	aatactaggg	3060
ttctgtgagt	ttgattaaag	tacggtgata	aatataagct	atgtggtggt	ggggctatac	3120
tactgaatga	aaaatgactt	gaaattttct	gcaattgaaa	aataaacacg	ttgaaacata	3180
acatgcaaca	ggttcacgat	tctttattcc	tgggcaatgt	aggagaaggt	gtaagagttg	3240
gtagcaaaaag	tttcagtggg	gtattttcca	ctttccagg	accatgtaaa	agacatagag	3300
taagtgccta	cctcgctagt	ttctgtggat	tcactagtgc	cattaagtgt	aatggtaagt	3360
atcatagggt	tagttttatc	accatgcaag	taaacttgac	tgacaatgtt	attttttagca	3420
gtttgacttt	gggtttttgg	ataggctaga	aggttaggca	taaattccaac	tgcattttgtg	3480
tatggatttg	cattagttag	gttcccattt	ctaagtttcc	agtaatgttt	tttaagttag	3540
gagttctcca	ttagaacacc	gttttggtca	aatctaagg	atatactaac	acttgcaacg	3600
gtgcctgtca	tggatgaaag	atctccagat	acagccaaag	cagctacagt	agctagtact	3660
tgactcccac	attttgtaag	aaccaaagta	aatttgcaat	cattatctga	atgaattctg	3720
cagttaggag	atgggtctgg	ggttgtccac	agggttaagt	tgtcatcatt	tttgtttctc	3780
attgtaatgg	cccctgagtt	gtcaaaagct	aaacccgctc	caagtttagt	aatcatggca	3840
ccgttttcat	tgtaatcaat	gccagagcca	attttagttt	ttattgggtt	gatatactgga	3900
gactcagatg	tgtttgtatc	aaactccaga	ccctttcctg	catttatagc	tatggcagta	3960
ttatcaaagt	ttagtccact	ggattttttt	atgctaactt	ccagtttttt	agtattgttt	4020
gatgcattaa	aaaggtatag	gcctctgtta	tagtttatgt	ccaagttatg	agatgcatta	4080
atatacaggg	gtccctgccc	cagtttaaga	ctgagttttg	tttgagcatc	aaatgggtaa	4140
ttccatctca	gaattaacaa	gttggtattt	atacgcatgc	caccgcccgt	tttaatttcc	4200
atgttgtttg	atgaatcata	accaatagct	cctgcaactt	tggttctaag	ggagttttgt	4260
tcaacggtga	cacctggtcc	agtaactact	gttagtgtat	cggagttttg	tgctacttgc	4320
aaaggaccgc	ttattttaat	tcctattttt	ccattattta	cataaatagg	atcttccatg	4380
ttaatgcccc	agctaccgtg	ggcagtagtt	agcgggggtg	atgcagttac	agtaaggggtg	4440
tcgctgtcac	tgccagagag	gggggctgat	gtttgcaggg	ctagctttcc	atctgacact	4500
gtaatggggc	ctttagtagc	aatgcttagt	ttggagtctt	gcacggtcag	tggggcttgt	4560
gactgtacgc	taagagcgcc	gctagtaact	atcagaggag	cgggtggtgc	cactgttagg	4620
gcgctgagg	taattgtaag	tgggtcggag	gtgtccaaac	ttatgtttga	ctttgttttt	4680
ttaagtggct	gagtaacagt	ggttacattt	tgggaggtga	ggtttccggc	cttgcctagg	4740
gtaagaccgc	tgcccatttt	aagcgcaagc	atgccgtggg	aggtgtccaa	aggttcggag	4800
acgcgtagag	agagaactcc	agggggactt	tcttggaac	cattgggtga	aacaaatgga	4860
ggggtaagaa	agggcacagt	tggaggcccg	gtttctgtgt	catatggata	cacgggggtg	4920
aaggtgtctt	cagacggtct	ggcgcttttc	atctgcaaca	atatgaagat	agtgggtgct	4980
gagggacaag	acatgagga	atttgacatc	ccatttaaac	tttgagaaaa	gtttgcagct	5040
aaaaggcggc	tgagatacca	gagttgggag	gaaggaaaag	aggtgatgct	gaataagctg	5100
gacaaagatt	tgctgactga	ttttaagtaa	gtaattttatt	cagtcgtagc	cgtccgccga	5160

gtctttcacc	gcgtaaaagt	tgggaataaaa	ctggtccggg	tagtggccgg	gaggtccaga	5220
aaaggggttg	aagtaaaccg	aaggcacgaa	ctcctcaata	aattgtagag	ttccaatgcc	5280
tccggagcgc	ggctccgagg	acgaggtctg	cagagttagg	atcgccctgac	ggggcgtaaa	5340
tgaagagcgg	ccagcgccgc	cgatctgaaa	tgtcccgctcc	ggacggagac	caagagagga	5400
gctcaccgac	tcgtcgttga	gctgaataacc	tcgccctctg	attttcagggt	gagttatacc	5460
ctgcccgggc	gaccgcaccc	tgtgacgaaa	gccgcccgcga	agctgcgccc	ctgagttagt	5520
catctgaact	tcggcctggg	cgtctctggg	aagtaccaca	gtggtgggag	cgggactttc	5580
ctggtacacc	agggcagcgg	gccaactacg	gggattaaag	ttattacgag	gtgtggtggt	5640
aatagccgcc	tgttcgagga	gaattcgggt	tcggtgggcg	cggattccgt	tgaccgggga	5700
tatcatgtgg	ggccccgcgc	tcagttagtt	tattcgggtt	gagtagtctt	gggcagctcc	5760
agccgcaagt	ccatttgtg	ctgggttaact	ccacatgtag	ggcgtgggaa	tttccctgct	5820
cataatggcg	ctgacgacag	gtgctggcgc	cgggtgtggc	cgctggagat	gacgtagttt	5880
tcgcgcttaa	atttgagaaa	gggcgcgaaa	ctagtcctta	agagtcagcg	cgagtatatt	5940
gctgaagaga	gcctccgcgt	cttcacagcg	gcgcggaagc	tgatcttcgc	ttttgtgata	6000
caggcagctg	cgggtgaggg	agcgcagaga	cctgtttttt	attttcagct	cttgttcttg	6060
gccctgtctt	tgttgaaata	tagcatacag	agtggaataa	atcctatttc	taagctcgcg	6120
ggtcgatacg	ggttcggttg	gcgccagacg	cagcgctcct	cctcctgctg	ctgccgcccgc	6180
tgtggatttc	ttgggctttg	tcagagtctt	gctatccggt	cgcccttgct	tctgtgtgac	6240
cgctgctggt	gctgcgcgtg	ccgctgccgc	cgggtgcagta	ggggctgtag	agatgacggt	6300
agtaaatgcag	gatgttacgg	gggaaggcca	cgccgtgatg	gtagagaaga	aagcggcggg	6360
cgaaggagat	gttgcccca	cagtcttgca	agcaagcaac	tatggcgctt	ttgtgccgcg	6420
gccacgagcg	gtagccttg	cgctgttggt	gctcttgggc	taacggcggc	ggctgcttag	6480
acttaccggc	cctggttcca	gtggtgtccc	atctacggtt	gggtcggcga	acaggcagtg	6540
ccggcggcgc	ctgaggagcg	gaggttgtag	cgatgctggg	aacggttgcc	aatttctggg	6600
gcgcggcgga	ggggaatgcg	accgaggggtg	acgggtgttc	gtctgacacc	tcttcggcct	6660
cggaaagcttc	gtctaggctg	tcccagctct	ccatcatctc	ctcctcctcg	tccaaaacct	6720
cctctgcttg	actgtcccag	tattcctcct	cgtccgtggg	tggcggcggc	ggcagctgca	6780
gcttcttttt	gggtgccatc	ctgggaagca	agggcccgcg	gctgctgata	gggctgcggc	6840
ggcgggggga	ttgggttgag	ctcctcgccg	gactgggggt	ccaggtaaac	cccccgctcc	6900
tttcgtagca	gaaactcttg	gcgggctttg	ttgatggctt	gcaattggcc	aaggatgttg	6960
ccctgggtaa	tgacgcaggc	ggtaagctcc	gcatttggcg	ggcgggattg	gtcttcgtag	7020
aacctaactc	cgtgggcgtg	gtagtctca	ggtacaaatt	tgcaaggta	agccgacgtc	7080
cacagccccg	gagtgagttt	caaccccgga	gccgcggact	tttcgctcag	cgagggacct	7140
tgacagctcaa	aggtaccgat	aatttgactt	tcgctaagca	gttgcgaaat	gcagaccagg	7200
gagcgggtcg	gggtgcatag	gttgacgca	cagtgacact	ccagtaggcc	gtcaccgctc	7260
acgtcttcca	tgatgtcgga	gtggtaggca	aggtagtggg	ctagctgcag	aagggtagcag	7320
tgaccccaaa	gcggcggagg	gcattcacgg	tacttaattg	gcacaaagtc	gctaggaagc	7380
gcacagcagg	tggcgggcag	aattcctgaa	cgctctagga	taaagttcct	aaagttttgc	7440
aacatgcttt	gactggtgaa	gtctggcaga	ccctgttgca	gggttttaag	caggcgcttcg	7500
gggaagataa	tgtccgccag	gtgcgcggcc	acggagcgct	cgttgaaggc	cgctccatagg	7560
tccttcaagt	tttgctttag	cagcttctgc	agctccttta	ggttgcgctc	ctccaggcat	7620
tgctgccaca	cgcccatggc	cgtttgccag	gtgtagcaca	gaaataagta	aacgcagctc	7680
cggacgtagt	cgccgcgcgc	ctcgcccttg	agcgtggaat	gaagcacgtt	ttgcccgagg	7740
cggttttcgt	gcaaaattcc	aaggtaggag	accaggttgc	agagctccac	gttggaattt	7800
ttgcaggcct	ggcgcacgta	gccctggcga	aaggtgtagt	gcaacgtttc	ctctagcttg	7860
cgctgcatct	ccgggtcagc	aaagaaccgc	tgcatgcact	caagctccac	ggtaacaagc	7920
actgcggcca	tcattagctt	gcgtcgctcc	tccaagtcgg	caggctcgcg	cgctcgaagc	7980
cagcgcgcca	gctgctcatc	gccaactcgc	ggtaggccct	cctcggtttg	ttcttgcaag	8040
tttgcatccc	tctccagggg	tcgtgcacgg	cgacgatca	gctcgctcat	gactgtgctc	8100
ataaccttgg	ggggtaggtt	aagtgcgggg	taggcaaagt	gggtgacctc	gatgctgctg	8160
ttcagcacgc	ctaggcgcgc	gttgtcacc	tcaagttcca	ccagcactcc	acagtgactt	8220
tcattttcgc	tgttttcttg	ttgcagagcg	tttgccgcgc	gtttctcgtc	gcgtccaaga	8280
ccctcaaaga	tttttgccac	ttcgctcagc	gaggcgatat	caggtagtac	agcgccctgc	8340
cgcaaggcca	gctgcttgct	cgctcggtg	cggttggcac	ggcaggatag	gggtatcttg	8400
cagtttttga	aaaagatgtg	ataggtggca	agcacctctg	gcacggcaaa	tacggggtag	8460
aagttgaggc	gcgggttggg	ctgcgatgtg	cggttttctt	ggcggttggg	gggtacgcgc	8520
ggtgagaaca	gtggcggttc	gtaggcaagg	ctgacatccg	ctatggcgag	gggcacatcg	8580
ctgcgctctt	gcaacgcgtc	gcagataatg	gcgcactggc	gctgcagatg	cttcaacagc	8640

acgtcgtctc	ccacatctag	gtagtcgcca	tgcccttggg	cccccgccc	gacttgttcc	8700
tcgtttgcct	ctgcgtcgtc	ctggctcttg	tttttatcct	ctgttgggtac	tgagcgatcc	8760
tcgtcgtctt	cgcttacaaa	acctgggtcc	tgctcgataa	tcacttcctc	ctcctcaagc	8820
gggggtgcct	cgacggggaa	ggtggtaggc	gcgttggcgg	catcggtgga	ggcgggtggtg	8880
gcgaactcaa	agggggcggt	taggctgtcc	tccttctcga	ctgactccat	gatctttttc	8940
tgccctatagg	agaaggaaat	ggccagtcgg	gaagaggagc	agcgcgaaac	cacccccgag	9000
cgcggacgcg	gtgcggcgcg	acgtccacca	accatggagg	acgtgtcgtc	cccgtcgccg	9060
tcgccgcccgc	ctccccgcgc	gccccaaaa	aageggctga	ggcggcgctc	cgagtccgag	9120
gacgaagaag	actcgtcaca	agatgcgctg	gtgccgcgca	cacccagccc	gcggccatcg	9180
acctcgacgg	cggatttggc	cattgcgtcc	aaaaagaaaa	agaagcgccc	ctctcccaag	9240
cccgaagcgc	cgccatcccc	agaggtgatc	tgggacagcg	aggaagaaag	agaagatgtg	9300
gcgctacaaa	tggtgggttt	cagcaacca	ccggtgctaa	tcaagcacgg	caagggaggt	9360
aagcgcacgg	tcgggcccgt	gaatgaagac	gacccagtgg	cgcggggtat	gcggacgcaa	9420
gaggaaaagg	aagagtccag	tgaagcggaa	agtgaagaca	cgggtgataaa	cccgttgagc	9480
ctgccgatcg	tgtctgcgtg	ggagaagggc	atggaggctg	cgcgcgcgtt	gatggacaag	9540
taccacgtgg	ataacgatct	aaaggcaaac	ttcaagctac	tgccctgacca	agtggaaagt	9600
ctggcggccg	tatgcaagac	ctggctaaac	gaggagcacc	gcgggttgca	gctgaccttc	9660
accagcaaca	agacctttgt	gacgatgatg	gggcgattcc	tcaggcgcta	cctgcagtcg	9720
tttgacagagg	taacctacaa	gcaccacgag	cccacgggct	gcgcgttgtg	gctgcaccgc	9780
tgccgtgaga	tcgaaggcga	gcttaagtgt	ctacacggga	gcattatgat	aaataaggag	9840
cacgtgattg	aaatggatgt	gacgagcgaa	aacgggcagc	gcgcgctgaa	ggagcactct	9900
agcaaggcca	agatcgtgaa	gaaccgggtg	ggccgaaatg	tggtgcagat	ctccaacacc	9960
gacgcaaggt	gctgcgtgca	tgacgcggcc	tgtccggcca	atcagttttc	cggaagtctc	10020
tgccgcatgt	tcttctctga	aggcgcaaa	gctcaggtgg	cttttaagca	gatcaaggct	10080
ttcatgcagg	cgctgtatcc	taacgcccag	accgggcacg	gtcaccttct	gatgccacta	10140
cggtgcgagt	gcaactcaaa	gcctgggcat	gcacctttt	tggaagggca	gctaccaaa	10200
ttgactccgt	tcgcccgtg	caacgcggag	gacctggacg	cggatctgat	ctccgacaag	10260
agcgtgctgg	ccagcgtgca	ccaccggcg	ctgatagtgt	tccagtgtgt	caaccctgtg	10320
tatcgcaact	cgcgcgcgca	gggcggaggc	cccaactgcg	acttcaagat	atcggcgccc	10380
gacctgctaa	acgcgttggg	gatggtgcgc	agcctgtgga	gtgaaaactt	caccgagctg	10440
cgcggatggg	ttgtgcctga	gtttaagtgg	agcactaaac	accagtatcg	caacgtgtcc	10500
ctgccagtgg	cgcatagcga	tgccgggcag	aacccctttg	atttttaaac	ggcgagacg	10560
gcaaggggtg	ggggtaaata	atcacccgag	agtgtacaaa	taaaaacatt	tgcttttatt	10620
gaaagtgtct	cctagtacat	tatttttaca	tgtttttcaa	gtgacaaaaa	gaagtggcgc	10680
tcctaactctg	cgcactgtgg	ctgcggaagt	agggcgagtg	gcgctccagg	aagctgtaga	10740
gctgttcctg	gattcgacgc	aggggtgggt	gtacctgggg	actgttaagc	atggagttgg	10800
gtaccccggt	ataaagggtc	atgggtgggt	tgtgatccat	gggagtttgg	ggccagttgg	10860
caaaggcgtg	gagaaacatg	cagcagaata	gtccacaggc	ggccgagttg	ggccctgca	10920
cgctttgggt	ggacttttcc	agcgttatac	agcggtcggg	ggaagaagca	atggcgctac	10980
ggcgcaggag	tgactcgtac	tcaaactggg	aaacctgctt	gagtcgttgg	tcagaaaagc	11040
caaagggctc	aaagaggtag	catgtttttg	agcgcgggtt	ccaggcaaa	gcatccaggt	11100
gtacgcccc	agtcctcgga	ccggccgtat	tgactatggc	gcaggcgagc	ttgtgtggag	11160
aaacaaagcc	tggaaagcgc	ttgtcatagg	tgcccaaaaa	atatggccca	caaccaagat	11220
ctttgacaat	ggcttttcagt	tcctgtctac	tggagcccat	ggcggcagct	gttgttgatg	11280
ttgcttgctt	cttttatggt	gtggcggtgc	cggccgagaa	gggcgtgcgc	aggtacacgg	11340
tctcgatgac	gcccgggtgc	ggctggtgca	cacggaccac	gtcaaagact	tcaaacaaaa	11400
cataaagaag	ggtgggctcg	tccatgggat	ccacctcaaa	agtcatgtct	agcgcgtggg	11460
cggagttggc	gtagagaagg	ttttggccca	ggtctgtgag	tgccgcccag	gacataaagt	11520
tactggagaa	tgggatgcgc	caaaggggtg	gatcgcaaa	aaactttttc	tgggtaatac	11580
tgtcaaccgc	ggtttttgcct	attagtgggt	agggcacgtt	ggcggggtaa	gctgtccctc	11640
cgcgcagtgt	gggagcgagg	tagcctacga	atcctgagtt	gttatgctgg	tgaagaattc	11700
caacctgctg	atactccttg	tatttagtat	cgtcaaccac	ttgcggctc	atgggctgga	11760
agtttctgaa	gaacgagtac	atgcggctct	tgtagctttc	tggaatgtag	aagccctggg	11820
agccaatat	gtagttggcc	aacatctgca	ccaggaacca	gtccttggtc	atgttgcaat	11880
gagctacgtt	gtagccctcc	ccgtcaactg	agcgtttaat	ctcaaaactc	ttgggagtaa	11940
gcaggcggtc	tgtgcccggc	cagctaacag	aagagtcaaa	ggtaatggcc	accttcttaa	12000
aggtgtgatt	aagatagaag	gttccgtcaa	ggtatgggat	ggagccagag	taggtgtagt	12060
aagggtcgta	gcctgatccc	aggggaaggg	tttcctttgt	cttcaagcgt	gtgaaggccc	12120

aaccgcgaaa	tgctgccag	ttgcgcgatg	ggatggagat	gggcacgttg	gtggcgttgg	12180
cgggatatggg	gtatagcatg	ttggcggcgg	aaaggtagtc	attaaaggac	tggtcgttgg	12240
tgatatttct	gagcatggct	tccagcgtgg	aggccgtgtt	gtgggccatg	gggaagaagg	12300
tgccgtaaag	acaaatgctg	tcaaacttaa	tgctagcccc	gtcaactcta	agatcgtttc	12360
ccagagagct	ctgcagaacc	atgttaacat	ccttcctgaa	gttccattca	tatgtatatg	12420
agcctggcag	gaggaggagg	tttttaattg	caaaaaactt	ttggggcacc	tgaatgtgaa	12480
agggcacgta	gcggccggtt	cccaacaaca	tggagcgata	acggaggccc	gcattgcggt	12540
ggtggttaaa	gggattaacg	ttgtccatgt	agtccagaga	ccagcgcgcc	ccaaggttaa	12600
tgtagcagtc	tacaagcccc	ggagccacca	ctcgcttggt	catgtagtcg	taggtgttgg	12660
ggttgtcaga	tatttccaca	ttggtggggt	tgtattttag	cttgtctggc	aggtacagcg	12720
caatatttga	gtaaaggaaa	tttctccata	ggttggcatt	taggttaatt	tccatggcaa	12780
agttgttacc	cactcctatt	tcattacgtg	ttgcaaaagt	ttcatctttt	gtccatgtag	12840
tatctccatt	atcgctgag	ccattgccat	tagccttaat	agcttgatag	gtgtcagtta	12900
ccccaatacc	cccaagagga	aaacaataat	ttggcaattc	atcctcagtt	ccatggtttt	12960
caatgattct	aacatctgga	ctacagctgt	ctacagcctg	attccacata	gaaaaatatc	13020
tggttctatc	acctaaggaa	tcaagcaaga	ggtgatagga	cagctctgtg	tttctgtctt	13080
gcaaatctac	cacggcattt	agctgcgatg	cctgaccagc	aagaacaccc	atggtgccag	13140
tgctgttata	atatactagg	ccaataaaat	tgtccctgaa	agcaatgtaa	ttgggtctgt	13200
ttggcataga	ttgttgaccc	aacatagctt	tagaattttc	atcacctttt	ccaggtttgt	13260
aagacagatg	tgtgtctggg	gtttccatat	ttacatcttc	actgtacaaa	accacttttg	13320
gttttagtagc	attgccttgc	cggctgttca	aagaggtagt	atttgagaag	aattgcaagt	13380
caaccttttg	aagaggcacc	cctttttcat	ccggaaccag	aacggattga	ccaccaaag	13440
gattttagtag	cctggcataa	gatccatagc	atggtttcat	gggagttgtt	tttttaagca	13500
ctctccctcc	tgccgcatta	gcatacagctt	cgttccactg	agattcgcca	atttgaggtt	13560
ctggttgata	ggaaggatct	gcgtatacag	gtttagcttg	tgtttctgca	ttgtctgac	13620
ctattttagtag	cccgcctttt	gtaattgttt	ctccagacaa	aggagcctgg	gcatagacat	13680
gtgttttctt	agtagcctga	tctcgagcgt	tttgcctctc	ttcttctctc	tcttcatctt	13740
catcttctct	ttcttcatcc	tccgcaactg	cccggccgct	atcttcggtt	tgttccact	13800
cacaggagtt	aggagcgccc	ttgggagcta	gagcgttgta	ggcagtgccg	gagtagggct	13860
taaaagtagg	ccccctgtcc	agcacgccgc	ggatgtcaaa	gtacgtggaa	gccatatcaa	13920
gcacacgggt	gtcaccacaa	gccagggtga	accgcgcttt	gtacgagtac	gcggtatcct	13980
cgcggtccac	agggatgaac	cgcagcgtca	aacgctggga	ccggtctgtg	gttacgtcgt	14040
gcgtaggtgc	caccgtgggg	tttctaaact	tgttattcag	gctgaagtac	gtctcgggtg	14100
cgcgggcaaa	ctgcaccagc	ccggggctca	ggtactccga	ggcgctcctg	cccagagatg	14160
gcataagaa	ccactgcggc	atcatcgaag	gggtagccat	cttgaaagc	gggcgcacgg	14220
cggctcagca	gtcctctctg	cggcgacatg	gacgcataca	tgacacatac	gacacgttag	14280
ctatttagaa	gcatacgtcg	cgttccaggg	attgcacccc	cagacccacg	atgctgttca	14340
gtgtgctttg	ccagtgtcca	ctggctacgg	gccgcacgca	tccgagaccg	ctggcggcac	14400
ggcgcaggga	cgcgcggcta	ggcggggtta	caacaacggc	ggacggccct	ggcagcacag	14460
gtttctgctg	ggtgtcagcg	gggggaggca	ggtccagcgt	tacaggtgtg	tgttgcccca	14520
gcactccggt	agccatgggc	gcgatgggac	gggtgggtgg	caggccttgc	tttagtgcct	14580
cctcgtacga	gggaggtcca	tctatttgcg	tcaccagagt	ttcttccctg	tccggcccg	14640
gacgcttttc	gccacgcccc	tctggagaca	ctgtctccac	ggcgggtgga	ggctcctcta	14700
cgggaggggc	gggatcaagc	ttactgttaa	tcttattttg	cactgcctgg	ttggccaggt	14760
ccaccacccc	gctaatagcca	gaggccaggc	catctaccac	cttttgttgg	aaatttttgt	14820
ctttcaactt	gtccctcagc	atctggcctg	tgtgctgttt	ccaggccttg	ctgccatagt	14880
tcttaattgg	ggaaccgaaa	tttttaattg	cgtccacag	cagaccccag	ctgaaggcgc	14940
caccgctcat	attgctggtg	ccgatattct	gccagtttcc	catgaacggg	cgcgagccgt	15000
gtcgcggggc	cagagacgca	aagttgatgt	cttccattct	acaaaatagt	tacaggacca	15060
agcagcgtg	agactccaga	ctttttattt	tgaattttcc	acatgcaact	tgtttttaat	15120
cagtgtctct	gcgcctgcaa	ggccacggat	gcaattccgg	gcacggcgcc	aatcgccgcg	15180
gcgatcagtg	gaataaggag	gggcaggata	ccgcgcgca	tgcgacggtg	cgcgcgcgcg	15240
cgcgcgggt	ggtgcgcacg	acgcattgcc	cccgtcaggc	cgtggccggc	catgcccctc	15300
ctacggtgca	ttcttctctg	gaatcccggc	accgggaaac	ggaggcggca	ggtgagggcc	15360
atatctgcaa	gaaccacaaa	gaccggcttt	taaacgatgc	tggggtggtg	gcgcgctgtt	15420
ggcagcacca	gggtcctgcc	tccttcgcga	gccaccctgc	gcacggaaat	cggggccagc	15480
acgggctggc	gacggcgacg	gcggcgcgcg	gttccagttg	tggttcggcg	tccggtagtc	15540
gctcgtcttc	tggggcggtg	ggtgtagcca	cgatagccgg	gggtaggcgc	gatggaagga	15600

tgtagggcat	attcgggcag	tagtgcgctg	gcggtgccgt	acttcctgga	acggcgcggg	15660
cgccgggggg	ctgaaacgcg	aaacatccac	gggtccgttt	gcacctccgt	agagggtttt	15720
gacggggccg	cagcgggccg	ctgcaccgcg	gcatctgcc	ccgcccaggc	aaccggggac	15780
gtttgtgtct	ccatgccctc	tgtggcagtg	gcaatactag	tgctactggg	ggtgggtatc	15840
tgaacgtcca	cggctgcac	gcccagtcct	ggtgccacct	gcttgattgg	ccgcacgcgg	15900
acctcgggct	ccagcccagg	ctccacggtc	attttttcca	agacatcttc	cagtcgctgg	15960
cgcttgggta	ccatcagctg	cacgggtggg	gccaagtac	cagactcgcg	ctttaggccg	16020
cgcttttctt	cggacggtgc	aagcgtgggc	agcacctgct	gcagtgtcac	gggcttttag	16080
ctaggtgttg	ggttgccctc	gtccagcggc	aacgcccaaca	tgctccttat	ccgctttccg	16140
taggcaaact	ccccgaggcg	ctcgttggcc	tgctcaagca	ggtcctcgtc	gccgtacacc	16200
tcatacata	cgcgcttgta	ggtgcgggtg	gagcgctcac	cgggcgtaaa	aactacgggtg	16260
gtgcccgggtc	gcaaaacacg	tcttacgctg	cgacctttcc	actgtaccgg	ccgcctgggc	16320
gcggttgcg	gcagcagttc	cacctcgctg	tcaagttcat	catcatcatc	atcttttctt	16380
ttctttttga	cccgttttag	ctttcggggc	ttgtaatcct	gctcttcctt	cttcgggggg	16440
ccatagatct	ccggcgcgat	gacctggagc	atctcttctt	tgattttgcg	cttgacata	16500
gcttcgttgc	gcgccgcgc	cgctggatc	atacaacagt	acgagtctaa	gtagtttttt	16560
cttgcaatct	agttgcgcg	ggggcggggtg	cgacggggca	cgcgaggcc	gctaaccgag	16620
tcgcgcaccc	agtacacggt	gccccgtcga	ccctgagtc	tagcactaat	ggccgcggct	16680
gctcgggcg	ccgctcgctg	cctggacctg	gggggcacag	tgacaatacc	cgcgccagc	16740
cttcgagcgg	cccgcattgg	cgcccgctcg	cggtgcgac	gtgcgcgggt	aagcagggcc	16800
gccgcgcgc	gttggggcgc	agtgcgggtg	cgggcgcggt	ggcgacgtgc	tacggccctc	16860
cgccgtctct	tcatttttag	ataacgcggg	gtcccgcgca	ccacggtctg	aatggccgcg	16920
tccactgtgg	acactggtgg	cgccgtgggc	gtgtagtgtg	gcgcctcttc	caccaccgcg	16980
tcaatggcgt	catcgacggt	ggtgcgcca	gtgcggccgc	gtttgtgcgc	gccccagggc	17040
gcgcggtagt	gccccgcac	gcgcactggg	tggtggtcgg	agcgcttctt	tgccccgcca	17100
aacatcttgc	ttgggaagcg	caggccccag	ctgtgttat	tgctggcgca	tataaggatg	17160
gacatgtttg	ctcaaaaagt	gcggctcgat	aggacgcgcg	gcgagactat	gccccgggc	17220
ttgtaaacgt	aggggcaggt	gcggcgctctg	gcgtcagtaa	tggtcactcg	ctggactcct	17280
ccgatgctgt	tgcgacgcg	tagcgtcccg	tgatctgtga	gagcaggaa	gttttactg	17340
acggtggtga	tggtgggggc	tgggcgggcg	gcaaaaatct	ggttctcggg	aaagcgattg	17400
aacacgtggg	tcagagaggt	aaactggcgg	atgagctggg	agtagacggc	ctggtcggtg	17460
tagaagctct	tggagtgcac	gggcaacagc	tcggcgccca	ccaccggaaa	gttgctgac	17520
tggtcgtgg	agcggaaggt	cacggggtct	tgcatcatgt	ctggcaacga	ccagtagacc	17580
tgctccgagc	cgcaggttac	gtcaggagtg	caaaggaggg	tccatgagcg	gatcccggtc	17640
tgagggtcgc	cgtagtgtga	tgcaaggtag	cagctgcggg	actgggtgaa	ggtgctgtca	17700
ttgcttatta	ggttgtaact	gcgtttcttg	ctgtcctctg	tcagggggtt	gatcacgggt	17760
ttcttctgag	gcttctcgac	ctcggttggt	gcagcggggg	cggcagcttc	tgccgctgcc	17820
tcggcctcag	cgcgcttctc	ctccgcccgt	gtggcaaaagg	tgctgcgcgc	aatggcatga	17880
tcgttcatgt	cctccaccgg	ctgcattgcc	gcggctgccc	cgttgaggtt	ctcttccgcg	17940
ccgctgccac	tggtgttgcc	gcgcctcg	ccatccccgc	cctgttcggg	gtcatctttt	18000
aagcttgcc	ggtaggcggt	cacatccaac	agtgcgggaa	tgttaccacc	ctccagggtc	18060
tcgtagggtga	tcctaaagcc	ctcctggaag	gtgtgcccgt	tgcggtatgc	caacaagttg	18120
ctcaggcggc	tgtaggtgaa	gtccaccccg	cactctggca	gcaaaatgat	gtctggatgg	18180
aaggcttcgt	ttgtatatac	cccaggcatg	acaagaccag	tgactgggtc	aaacccaggt	18240
ctgaagtgtc	gggtgtcaaa	ctttaccccg	atgtcgcttt	ccagaacccc	gttctgcctg	18300
cccactttca	agtagtgctc	cacgatcgcg	ttgttcataa	ggtctatggt	catggtctcg	18360
gagtagttgc	cctcgggcag	cgtgaactcc	accactcat	atttcagctc	cacctgtttg	18420
tccttagtaa	gcgagcgca	caccatcacc	cgcgcttaa	acttattggt	aaacatgaac	18480
tcgttcacat	ttggcatgtt	ggtatgcagg	atggttttca	ggtcgccgcc	ccagtgcgaa	18540
cggtcgtcaa	gattgatggt	ctgtgtgctt	gcctcccccg	ggctgtagtc	attgttttga	18600
atgaccgtgg	ttagaaagtt	gctgtggtcg	ttctggtagt	tcagggatgc	cacatccgtt	18660
gacttggtgt	ccacaaggta	cacacgggtg	gtgtcgaata	ggggtgccaa	ctcagagtaa	18720
cggtgctgt	ttctcccccc	ggtagggcgc	aggtaccgcg	gaggcacaaa	cgggcggtcc	18780
aggggagcat	cgaaggggga	acccagcgcc	gccgccactg	gcgcgcgcgt	caccacgctc	18840
tcgtaggagg	gaggaggacc	ttcctcatat	atcgccgcgc	gctgcatact	aaggggaata	18900
caagaaaacc	aacgctcggt	gccatggcct	tggtgagttt	tttattttgc	atcatgcttt	18960
tttttttttt	ttttaaaca	cttccccag	cctggggcga	agggtgcgaa	acgggttgcc	19020
actccctccc	aaatccagga	cgctgctgtc	gtctgccgag	tcactgctct	cccacaccag	19080

accccgctga	cggctcgtgcc	tttgacgacg	ggtggggcggg	cgcggggcggg	gcacatccct	19140
gtgctcctgc	gcatacgtct	tccatctact	catcttgtcc	actaggctct	ctatcccgtt	19200
gttgggaaat	gccggaggca	ggttcttttc	gcgctgcggc	tgcagcagcg	agttgttttag	19260
gtactcctcc	tgcgccagca	ggcgcgggcg	ggtggtgcga	gtgctggtaa	aagaccctat	19320
caagcttgga	aatgggctac	tgcacatctga	ccgcggggcc	gcagcgccta	gatcggacaa	19380
gctgcttggc	ctgcggaagc	tttcctttcg	cagcgccgcc	tctgcctgct	cgcgctgttg	19440
caactctagc	agggtctgcg	ggtgcgggga	aaacacgctg	tcgtctatgt	cgtcccagag	19500
gaatccatcg	ttaccctcgg	gcacctcaaa	tcccccggtg	tagaaaccag	ggggcggtag	19560
ccagtgcggg	ttcaagatgg	cattggtgaa	atactcgggg	ttcacggcgg	ccgcgcgatg	19620
caagtagtcc	attaggcgat	tgataaacgg	ccggttttag	gcatacatgc	ccggttccat	19680
gttgcgcgcg	gtcatgtcca	gcgccacgct	ggcgcttacc	ccgtcgcgca	tcagggttaag	19740
gttcacgctc	tgctgcacat	agcgcaagat	gcgctcctcc	tcgctgttta	aactgtgcaa	19800
cgaggggatc	ttctgccgcc	ggttggtcag	caggtagttc	agggttgcct	ccaggctgcc	19860
cgtgtcctcc	tgccccagcg	cgcggctgac	acttgtaatc	tcctggaaag	tatgctcgtc	19920
cacatgcgcc	tgacctatgg	cctcgcggta	cagtgtcagc	aagtgcaccta	ggtatgtgtc	19980
ccggggacacg	ctgccactgt	ccgtgaaggg	cgctatttagc	agcagcaaca	ggcgcgagtt	20040
gggcgtcagc	aagctagaca	cggtcgcgcg	gtcgctgtg	ggagcccgca	ccccccacag	20100
cccctgcaag	ttcttgaaag	cctggctcag	gtttacggtc	tgcaggcctt	gtctactggg	20160
ctggaaaaaa	tagtctggcc	cggactggta	cacctcactt	tgcgggtgtct	cagtcaccat	20220
tagccgcagt	gcgctcacaa	agttgggtga	gtcctcctgt	ccccgcggca	cgttggcggg	20280
ctgtgtactc	aggaaggcgt	ttagtgaac	catggagccc	aggttgccct	gctgtcgcg	20340
gcgctcacgc	tgccgcacgg	ctcgcgcac	atccccacc	agccggtcca	ggttggctcg	20400
cacgttgccg	ctggtgtaac	gagccacgcg	ctgaagcagc	gcgtcgtaga	ccaggccggc	20460
ctcatcgggc	cggatggccc	tgttttcggc	cagcgcgttt	acgatcgcca	gcaccttctc	20520
gtgctggggg	tttgcgcgcg	ccgggaccac	cgcttccaga	attgcgga	gccggttggc	20580
ctgcggctgc	tgccggaacg	cgtcagggtt	acgcgcagtc	agcgacatga	tgcgggtccat	20640
gacctggcgc	cagtcgtccg	tggagttaag	gcgggacggc	tggctctgca	gcgcgcgccg	20700
caccgcccggg	tccgttgctg	cttgcatcat	ctgatcagaa	acatcacgcg	ttagtactcg	20760
ccgtcctctg	gctcgtactc	atcgctcctg	tcataattcct	ccacgcgcgc	gacgttgcca	20820
gcgcgcgcgg	gtgccaccgc	cagcccaggt	ccggccccag	ctgcctccag	ggcgcgctcg	20880
cttggggccc	agcgcaggtc	agcgcccgcg	tcaaagtagg	actcggcctc	tctatcgccg	20940
ctgcccgtgc	cagccagggc	cctttgcagg	ctgtgcatca	gctcgcggtc	gctgagctcg	21000
cgccgcgcgg	tcacgctcac	ggccttgctg	atgcgctcgt	tgcgataaac	gcccaggctg	21060
tcgctcaagg	taagcacctt	caacgccatg	cgcatgtaga	acccctcgat	ctttacctcc	21120
ttgtctatgg	gaacgtaagg	ggtatggtat	atcttgcggg	cgtaaaactt	gcccagactg	21180
agcatggaat	agttaatggc	ggccaccttg	tcagccaggc	tcaagctgcy	ctcctgcacc	21240
actatgctct	gcagaatgtt	tatcaaatcg	agcagccagc	ggccctcggg	ctctactatg	21300
tttagcagcg	catccctgaa	tgcctcgttg	tccttgctgt	gctgcactat	aaggaaacgc	21360
tgcgccatga	gcggcttgct	atgttggttt	tgctccagcg	cgcttacaaa	gtcccacaga	21420
tgcatcagtc	ctatagccac	ctcctcgcgc	gccacaagcg	tgcgcacgtg	gttggttaaag	21480
cttttttgaa	agttaatctc	ctggttcacc	gtctgtctgt	acgcgggttac	caggctcggc	21540
gccgccacgt	gtgcgcgcgc	gggactaatc	ccggtcgcgc	cgctcgggtc	aaagtccctc	21600
tcgcgcagca	accgctcgcg	gttcaggcca	tgcgcgaact	cgcgccctgc	gtggaacttt	21660
cgatcccgcg	tctcctcggg	ctcctctccc	tcgcggctgc	gaaacagggt	ctgcgcgggc	21720
acgtacgcct	cgcgcgtgtc	acgcttcagc	tgcacccttg	ggtgtcgctc	aggagagggc	21780
gctcctagcc	gcgccaggcc	ctcgcctctc	tccaagtcca	ggtagtgcgc	ggcccggcgc	21840
cgcggggggt	cgtaatcacc	atctgcgcgc	cggtcagccg	cggatgttgc	ccctcctgac	21900
gcggtaggag	aaggggaggg	tgccttgcat	gtctgcgcgt	gctcttgctc	ttgccgctgc	21960
tgaggagggg	ggcgcatctg	ccgcagcacc	ggatgcatct	gggaaaagca	aaaaaggggc	22020
tcgtccctgt	ttccggagga	atgtgcaagc	ggggtcttgc	atgacgggga	ggcaaacccc	22080
cgttcgcgcg	agtccggccg	gcccagagat	cgaaccgggg	gtcctgcgac	tcaacccttg	22140
gaaaataacc	ctccggctac	agggagcgag	ccacttaatg	ctttcgcttt	ccagcctaac	22200
cgcttacgcc	gcgcgcggcc	agtgcccaaa	aaagctagcg	cagcagccgc	cgcgcttggg	22260
aggaagccaa	aaggagcgct	cccccgttgt	ctgacgtcgc	acacctgggt	tcgacacgcg	22320
ggcggttaacc	gcattggatca	cggcgagcgg	ccggatccgg	ggttcgaacc	ccggtcgctc	22380
gccatgatac	ccttgccaat	ttatccacca	gaccacggaa	gagtgcgccg	ttacaggctc	22440
tccttttgca	cggctctagag	cgtcaacgac	tgcgcacgcc	tcaccggcca	gagcgtcccg	22500
accatggagc	actttttgcc	gctgcgcaac	atctggaacc	gcgtccgcga	ctttccgcgc	22560

gcctccacca	cgcgcgcgcg	catcacctgg	atgtccaggt	acatctacgg	atatcatcgc	22620
cttatgttgg	aagacctcgc	ccccggagcc	ccggccaccc	tacgtctggc	cctctaccgc	22680
cagccgcgcg	cgcacttttt	ggtgggatat	cagtacctgg	tgcggacttg	caacgactac	22740
gtctttgact	caagggctta	ctcgcgtctc	aggtaacacc	agctctcgcg	gccgggtcac	22800
cagaccgtta	actgggtccg	tatggccaac	tgcacttaca	ccatcaaac	gggcgcatac	22860
caccgctttg	tggacatgga	tgacttccag	tctacctca	cgcaggtgca	gcaggccata	22920
ttagccgagc	gcgttgtcgc	cgacctggcc	ctgcttcagc	cgatgagggg	cttcggggtc	22980
acacgcattg	gaggaagagg	gcgccacct	cgcccaact	ccgcgcgcgc	cgtagcgata	23040
gatgcaagag	atgcaggaca	agaggaagga	gaagaagaag	tgcgcgtaga	aaggctcatg	23100
caagactact	acaaagacct	gcgcgatgt	caaaacgaag	cctggggcat	ggccgaccgc	23160
ctgcgcattc	agcaggccgg	acccaaggac	atggtgcttc	tgtcgacct	ccgcgctc	23220
aagaccgcct	actttaatta	catcatcagc	agcacctccg	ccagaaacaa	ccccgaccgc	23280
caccgcgtgc	cgcccgccac	ggtgctcagc	ctaccttgcg	actgtgactg	gttagacgcc	23340
tttctcgaga	ggtttttccg	tccggtcgat	gcggactcgc	tcaggctccct	cggtggcgga	23400
gtacctacac	aacaattgtt	gagatgcac	gttagcgccg	tatccctgcc	gcacggcagc	23460
ccccgcgcaa	cccataaccg	ggacatgacg	ggcggcgctc	tccaactgcg	ccccgcgag	23520
aacggccgcg	gcctaccgga	gacctatgcg	gcgcgcgcg	gggagatgat	cgagcgcttt	23580
gtcgaccgcc	tcccggtgcg	ccgtcgtcgc	cgcggtgtcc	cccccccc	accgcgcgca	23640
gaagaagaag	aagaagggga	ggcccttatg	gaagaggaga	ttgaagaaga	agaggcccct	23700
gtagcctttg	agcgcgaggt	gcgcgacact	gtcgcgcgagc	tcacccgtct	cttgaggag	23760
gagttaaccg	tgtcggcgcg	caactcccag	tttttcaact	tcgcggtgga	cttctacgag	23820
gccatggagc	gccttgaggc	cttgggggat	atcaacgaat	ccacgttgcg	acgtctgggt	23880
atgtacttct	tctgtggcaga	acacaccgcc	accacctca	actacctctt	tcagcgcttg	23940
cgaactacg	ccgtcttcgc	ccggcacgtg	gagctcaatc	tcgcgcaggt	ggtcatgcgc	24000
gcccgcgatg	ccgaaggggg	cgtggtctac	agccgcgtct	ggaacgaggg	aggcctcaac	24060
gccttctcgc	agctcatggc	ccgcctctcc	aacgacctcg	ccgccaccgt	ggagcgagcc	24120
ggacgcggag	atctccagga	ggaagagatc	gagcagttca	tggccgaaat	cgctatcaa	24180
gacaactcag	gagacgtgca	ggagattttg	cgccaggccg	ccgtcaacga	caccgaaatt	24240
gattctgtcg	aactctcttt	caggttcaag	ctcaccgggc	ccgtcgtctt	cacgcagagg	24300
cgccagattc	aggagatcaa	ccgcgcgctc	gtcgcgttcg	ccagcaacct	ccgcgcgcag	24360
caccagctcc	tgcgcgcgcg	cggcgccgac	gtgcccctgc	cccctctccc	ggcgggtccc	24420
gagccccccc	tacctccggg	ggcccgcgcg	cgtaaccgct	tttagatgca	tcacccaagg	24480
acacccccgc	ggcccaccgc	ccgcgcgcgc	gtaccgtagt	cgcgccgcgc	ggatgcccgc	24540
tcttgcaagt	catcgacgcc	gccaccaacc	agcccctgga	aatcagggtat	caactggacc	24600
tagccccgcg	cctgaccgcg	ctatgcgagg	taaacctgca	ggagctcccg	cctgacctgt	24660
cgccgcggga	gctccagacc	atggacagct	cccatctgcg	cgatgttgtc	atcaagctcc	24720
gaccgcgcgc	cgcgacatc	tggactttgg	gctcgcgcgc	cgtggtgggt	cgatccacca	24780
taactcccct	cgagcagcca	gacggtcaag	gacaagcagc	cgaagtagaa	gaccaccagc	24840
caaaccgcgc	aggcgagggg	ctcaaattcc	cactctgctt	ccttggtgcg	ggtcgtcagg	24900
tcaacctcgt	gcaggatgta	cagcccgtgc	accgctgcc	gtactgcgca	cgtttttaca	24960
aaagccagca	cgagtgttcg	gcccgtcgca	gggacttcta	ctttcaccac	atcaacagcc	25020
actcctccaa	ctggtggcgc	gagatccagt	tcttcccgat	cggctcgcat	cctcgaccgc	25080
agcgtctctt	tgtcacctac	gatgtagaga	cctatacttg	gatggggggc	tttgggaagc	25140
agctcgtgcc	cttcatgctg	gttatgaagt	tcggcgagga	tgagcctctg	gtgaccgcgc	25200
cgcgagacct	agccgtggac	cttggatggg	accgctggga	acaagaccgc	cttaccttct	25260
actgcatcac	cccagaaaaa	atggccatag	gtcgccagtt	taggaacctt	cgcgaccacc	25320
tgcaaatgct	aatggcccgt	gacctgtgga	gctcattcgt	cgcttccaac	cctcatcttg	25380
cagactgggc	cctgtcagaa	cacgggctca	gctcccctga	ggagctcacc	tacgaggaac	25440
ttaaaaaatt	gccctccatc	aagggcaccc	cgcgcttctt	ggaactttac	atcgtggggc	25500
acaacatcaa	cggcttcgac	gagatcgtgc	tcgcgcgcca	ggtaatatac	aaccgttccg	25560
aggtgcccgg	acccttcgcg	atcacacgca	actttatgcc	tcgcgcggga	aagatacttt	25620
tcaacgatgt	caccttcgcc	ctgcccaccc	cgcgttccaa	aaagcgcacg	gactttttgc	25680
tctggggagca	gggcccgatgc	gacgacactg	acttcaataa	ccagtacctc	aaagtcatgg	25740
ttagggacac	ctttgcgctc	acccacacct	cgtctcggaa	ggccgcgcag	gcatacgccg	25800
tacccgtaga	aaagggatgc	tgcgcctacc	aggccgtcaa	ccagttctac	atgctaggct	25860
cttaccgctc	ggaggccgac	gggtttccga	tccaagagta	ctggaaagac	cgcaagaggt	25920
ttgtcctcaa	ccgcgagctg	tggaaaaaaa	agggacagga	taagtatgac	atcatcaagg	25980
aaacctggga	ctactgcgc	ctagacgtgc	aggtcaccgc	cgagctgggt	aacaagctgc	26040

gcgactccta	cgcctccttc	gtgcgtgacg	cggtaggtct	cacagacgcc	agcttcaacg	26100
tcttccagcg	tccaaccata	tcatecaact	cacatgccat	cttcaggcag	atagtcttcc	26160
gagcagagca	gcccccccg	agcaacctcg	gtcccgacct	cctcgctccc	tcgcacgaac	26220
tatacgattta	cgtgcgcgcc	agcatccgcg	gtggaagatg	ctaccctaca	tatcttggaa	26280
tactcagaga	gccccctctac	gtttacgaca	tttgcggcgt	gtacgcctcc	gcgctcacc	26340
accccatgcc	atgggggtccc	ccactcaacc	catacgagcg	cgcgcttgcc	gcccgcgcat	26400
ggcagcaggc	gctagacttg	caaggatgca	agatagacta	cttcgacgcg	cgctgtctgc	26460
ccgggggtctt	taccgtggac	gcagaccccc	cggacgagac	gcagctagaç	ccactaccgc	26520
cattctgttc	gcgcgaaggcg	ggcgcgctct	gctggaccaaa	cgagcgccca	cgcgagagag	26580
tagccaccag	cgttgacctt	gtcacccctgc	acaaccgcgg	ttggcgcggtg	cacctgggtgc	26640
ccgacgagcg	caccaccgtc	tttcccgaat	ggcgggtgct	tgcgcgcgaa	tacgtgcagc	26700
taaacatcgc	ggccaaggag	cgcgcgcatc	tgacacaaaa	ccaaaccctg	cgctccatcg	26760
ccaagttgct	gtccaacgcc	ctctacgggt	cgtttgccac	caagcttgac	aacaaaaaga	26820
ttgtcttttc	tgaccagatg	gacgcggcca	ccctcaaagg	catcaccgcg	ggccaggtga	26880
atatcaaate	ctcctcgttt	ttggaaactg	acaatcttag	cgcagaagtc	atgcccgtct	26940
ttgagaggga	gtactcaccc	caacagctgg	ccctgcgaga	cagcgatgcg	gaagagagtg	27000
aggacgaacg	cgccccacc	cccttttata	gcccccttc	aggaacaccc	ggtcacgtgg	27060
cctacacctta	taaaccaatc	accttccttg	atgccgaaga	ggcgacatg	tgtcttcaca	27120
ccctggagcg	agtggacccc	ctagtggaca	acgaccgcta	ccccctccac	ttagcctcct	27180
tcgtgctggc	ctggacgcga	gccttcgtct	cagagtgggtc	cgagtctcta	tacgaggagg	27240
accgcggaac	accgctcgag	gacaggcctc	tcaagtctgt	atacggggac	acggacagcc	27300
tttctgtcac	cgagcgtgga	caccggctca	tggaaaccag	aggtaagaaa	cgtatcaaaa	27360
agcatggggg	aaacctggtt	tttgaccccg	aacggccaga	gctcacctgg	ctcgtggaat	27420
gcgagaccgt	ctgcggggcc	tgccggcgcg	atgcctactc	cccgggaatcg	gtatttctcg	27480
cgcccaagct	ctacgccctt	aaaagtctgc	actgccccctc	gtgcggcgcc	tcctccaagg	27540
gcaagctgcg	cgccaaggcg	cacgcgcgcg	aggggctgga	ctatgacacc	atggtcaaat	27600
gctacctggc	cgacgcgcag	ggcgaagacc	ggcagcgctt	cagcaccagc	aggaccagcc	27660
tcaagcgcac	cctggccagc	gcgcagcccg	gagcgcaccc	cttcaccgtg	accagacta	27720
cgctgacgag	gacctgtcgc	ccgtggaaag	acatgaccct	ggcccgctctg	gacgagcacc	27780
gactactgcc	gtacagcgaa	agccgcacca	accgcgcaaa	cgaggagata	tgctggatcg	27840
agatgccgta	gagcacgtga	ccgagctgtg	ggaccgcctg	gaactgcttg	gtcaaacgct	27900
caaaagcatg	cctacggcgg	acggcctcaa	accgttgaaa	aactttgctt	ccttgcaaga	27960
actgctatcg	ctggggcgcg	agcgccttct	ggcgcatctg	gtcagggaaa	acatgcaagt	28020
cagggacatg	cttaacgaag	tggccccctt	gctcagggat	gacggcagct	gcagctctct	28080
taactaccag	ttgcagccgg	taataggtgt	gatttacggg	cccaccggct	gcggtaatg	28140
gcagctgctc	aggaacctgc	tttcttccca	gctgatctcc	cctaccccg	aaacggtttt	28200
cttactgcc	ccgcaggtag	acatgatccc	cccactgaa	ctcaaagcgt	gggaaatgca	28260
aatctgtgag	ggtaactacg	cccctgggccc	ggatggaacc	attataccgc	agtctggcac	28320
cctccgcccc	cgctttgtaa	aaatggccta	tgacgatctc	atcctggaac	acaactatga	28380
cgttagtgat	cccagaaata	tcttcgccca	ggccgcgcgc	cgtgggccc	ttgcatcat	28440
tatggacgaa	tgcatggaaa	atctcggagg	tcacaagggc	gtctccaagt	tcttccacgc	28500
atttcttct	aagctacatg	acaaatttcc	caagtgaccc	ggatacactg	tgctgggtgg	28560
tctgcacaac	atgaatcccc	ggagggatat	ggctgggaac	atagccaacc	taaaaaata	28620
gtccaagatg	catctcatat	ccccacgtat	gcacccatcc	cagcttaacc	gctttgtaaa	28680
cacttacacc	aagggcctgc	ccctggcaat	cagcttgcta	ctgaaagaca	tttttaggca	28740
ccacgcccag	cgctcctgct	acgactggat	catctacaac	accaccccg	agcatgaagc	28800
tctgcagtgg	tgctacctcc	accccagaga	cgggcttatg	cccatgtatc	tgaacatcca	28860
gagtcacctt	taccacgtcc	tggaaaaaat	acacaggacc	ctcaacgacc	gagaccgctg	28920
gtcccgggcc	taccgcgcgc	gcaaaacccc	taaataaaga	cagcaagaca	cttgcttgat	28980
caaaatccaa	acagagtctg	gtttttatct	atgtttttaa	ccgcattggg	aggggaggga	29040
gccttcaggg	cagaaacctg	ctggcgcgaga	tccaacagct	gctgagaaac	gacattaagt	29100
tcccgggtca	aagaatccaa	ttgtgccaac	agagccgtca	acttgctatc	gcgggaggat	29160
gaacgggaag	ctgcactgct	tgcaagcggg	ctcaggaaag	caaagtcagt	cacaatcccc	29220
cgggcggtgg	ctgcagcggc	tgaagcgggc	gcggaggctg	cagtctccaa	cggcgttcca	29280
gacacggtct	cgtaggtcaa	ggtagtagag	tttgcgggca	ggacggggcg	accatcaatg	29340
ctggagccca	tcacattctg	acgcaccccg	gcccattggg	gcatgcgcgt	tgtcaaatat	29400
gagctcacaa	tgcttccatc	aaacgagttg	gcgctcatgg	cggcggtcgc	tgcaaaacag	29460
atacaaaact	acatgagacc	cccaccttat	atattcttcc	ccacccttaa	gccccgccc	29520


```

tcgatggcaa acagctatta tgggtattat ggggtgctagc gacatgaggt tgccccgatat 29580
tcagtgtcgc tgatttgtat tgtctgaagt tgttttttacg ttaagttgat gcagatcaat 29640
taatacgata cctgcgtcat aattgattat ttgacgtggg ttgatggcct ccacgcacgt 29700
tgtgatatgt agatgataat cattatcact ttacgggtcc tttccgggtga tccgacaggt 29760
tacggggcgg cgacctcgcg ggttttcgct atttatgaaa attttcgggt ttaaggcggt 29820
tccgttcttc ttcgtcataa cttaatgttt ttatttataaa taccctctga aaagaaagga 29880
aacgacaggt gctgaaagcg aggcctttttg gcctctgtcg tttcctttct ctgtttttgt 29940
ccgtggaatg aacaatggaa gttaacggat ccaggccgcg agcaaaaggc cagcaaaagg 30000
ccaggaaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 30060
agcatcacia aaatcaacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 30120
accaggcggt tccccctgga agctccctcg tgcgtctcc tggtccgacc ctgccgctta 30180
ccggatacct gtccgccttt tcccttcggt gaagcgtggc gctttctcat agctcacgct 30240
gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc 30300
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtaa 30360
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 30420
taggcggtgc tacagagttc ttgaagtggg ggcttaacta cggctacact agaagaacag 30480
tatttggtat tccgctctg ccaaagccag ttacctcgg aaaaagagtt ggtagctctt 30540
gatccggcaa acaaaccacc gctggtagcg gtgggttttt tgtttgcaag cagcagatta 30600
cgcgagaaa aaaaggatct caagaagatc ctttgatctt ttctacggg tctgacgctc 30660
agtggaaacy aaactcacgt taagggattt tggctcatcag attatcaaaa aggatcttca 30720
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat cttaaagtata tatgagtaaa 30780
cttggtctga cagttaccaaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 30840
ttcgttcatc catagttgcc tgactccccg tagtgtagat aactacgata cgggaggggt 30900
taccatccgg cccagtgct gcaatgatac cgctgaccc acgctcaccg gctcctgatt 30960
tatcagcaat aaaccagcca gccggaagtg ccgagcgagc aagtggtcct gcaactttat 31020
ccgcctccat ccagctctatt agttgttgcc ggggaagctag agtaagtagt tcgccagtta 31080
atagttttcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg 31140
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt 31200
tgtgcaaaaa agcgggttagc tccttcggtc ctccgatagt tgtcagaagt aagttggccg 31260
cagtggtatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 31320
taagatgctt ttctgtgact ggtgagtatt caaccaagaa tacgggataa taccgcgcca 31380
catagcagaa ctttaaaagt gctcatcatt gggaaacgtt ctccggggcg aaaaactca 31440
aggatcttac cgctgttgag atccagttcg atgtaacca ctccgcgacc caagtgatct 31500
tctgcatctt ttactttcac cagcgtttct ggggtgagcaa aaacaggaag gcaaaatgcc 31560
gcaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactttt cctttttcaa 31620
tattattgaa gcatttatca gggttattgt ctcacagcg gatacatatt tg 31672

```

<210> 4

<211> 30365

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 4

```

gaatcgcca ggcgaatta actataacgg tcctaaggta gcgtcatcat cataatatac 60
cttattttgg attgaagcca atatgataat gaggggggtg agtttgtgac gtggcgcggg 120
gcgtgggaac gggcggggtg acgtagggtt tagggcgag taacttgcac gtattgggaa 180
ttgtagtttt tttaaaatgg gaagttacgt atcgtgggaa aacggaagtg aagatttgag 240
gaagttgtgg gttttttggc tttcgtttct gggcgtaggt tcgcgtgcgg ttttctgggt 300
gtttttttgt gactttaacc gttacgtcat tttttagtc tatatatact cgtctgtgac 360
ttggcccttt ttacactgtg actgattgag ctggtgccgt gtcgagtggt gttttttaat 420
agggtttttt actggtaagg ctgactgtta tggctgccgc tgtggaagcg ctgtatgttg 480
ttctggagcg ggagggtgct attttgccta ggataacttc gtataatgta tgctatacga 540
agttatggcg cgccagatct gtttgtcacg cccgcacctg gttttgcttc aggaaatatg 600

```

actacgtccg	gcggttcatt	tggcatgaca	ctacgaccaa	cacgatctcg	gttgtctcgg	660
cgcactccgt	acagtaggga	tgcctacct	ccttttgaga	cagagacccg	cgctaccata	720
ctggaggatc	atccgctgct	gcccgaatgt	aacactttga	caatgcacaa	cgtaggttac	780
gtgcgaggtc	ttccctgcag	tgtgggattt	acgctgattc	aggaatgggt	tgttccctgg	840
gatatggttc	tgacgcggga	ggagcttgta	atcctgagga	agtgtatgca	cggtgcctg	900
tgttggtcca	acattgatata	catgacgagc	atgatgatcc	atggttacga	gtcctgggct	960
ctccactgtc	attgttccag	tcccggttcc	ctgcagtgc	tagccggcgg	gcaggttttg	1020
gccagctggt	ttaggatggt	ggtggatggc	gccatgttta	atcagagggt	tatatggtac	1080
cgggagggtg	tgaattacaa	catgccaaaa	gaggtaatgt	ttatgtccag	cggttttatg	1140
aggggtcgcc	acttaatcta	cctgcgcttg	tggatgatg	gccacgtggg	ttctgtggtc	1200
cccgccatga	gctttggata	cagcgccctg	cactgtggga	ttttgaacaa	tattgtgggt	1260
ctgtgctgca	gttactgtgc	tgatttaagt	gagatcaggg	tgcgctgctg	tgcccgagg	1320
acaaggcgctc	tcatgctgcg	ggcgggtgcga	atcatcgctg	aggagaccac	tgccatgttg	1380
tattcctgca	ggacggagcg	gcggcggcag	cagtttattc	gcgcgctgct	gcagcaccac	1440
cgccctatcc	tgatgcacga	ttatgactct	acccccatgt	aggcgtggac	ttcccccttcg	1500
ccgcccgttg	agcaaccgca	agttggacag	cagcctgtgg	ctcagcagct	ggacagcgac	1560
atgaacttaa	cggagctgcc	cggggagttt	attaatatca	ctgatgagcg	tttggtcga	1620
caggaaaccg	tgtggaatat	aacacctaag	aatatgtctg	ttacccatga	tatgatgctt	1680
tttaaggcca	gccggggaga	aaggactgtg	tactctgtgt	gttgggaggg	aggtggcagg	1740
ttegaacta	gggttctgtg	agtttgatta	aggtacggtg	atcaatataa	gctatgtggt	1800
ggtggggcta	tactactgaa	tgaaaaatga	cttgaaattt	tctgcaattg	aaaaataaac	1860
acgttgaaac	ataacatgca	acaggttcac	gattctttat	tcctgggcaa	tgtaggagaa	1920
ggtgtaagag	ttggtagcaa	aagtttcagt	ggtgtatttt	ccactttccc	aggaccatgt	1980
aaaagacata	gagtaagtgc	ttacctcgct	agtttctgtg	gattcactag	tgccattaag	2040
tgtaatggta	agtatcatag	gttttagttt	atcaccatgc	aagtaaactt	gactgacaa	2100
gttattttta	gcagtttgac	tttgggtttt	tggataggct	agaaggttag	gcataaatcc	2160
aactgcattt	gtgtatggat	ttgcattagt	tgagttccca	tttctaaagt	tcagtaaatg	2220
ttttttaagt	gaggagtctt	ccattagaac	accgttttgg	tcaaactctaa	ggaataact	2280
aacacttgca	acggtgcctg	tcattggatga	aagatctcca	gatacagcca	aagcagctac	2340
agtagctagt	acttgactcc	cacattttgt	aagaaccaa	gtaaatttgc	agtcattatc	2400
tgaatgaatt	ctgcagttag	gagatgggtc	tgggtgtgtc	cacagggtaa	gtttgtcatc	2460
atttttgttt	octattgtaa	tggcccctga	gttgtaaaag	cttaaacccg	ctccaagttt	2520
agtaatcatg	gcaccgtttt	cattgtaatc	aatgccagag	ccaatttttag	tttttattgg	2580
gttgatctct	ggagactcag	atgtgtttgt	atcaaactcc	agaccctttc	ctgcatttat	2640
agctatggca	gtattatcaa	agtttagtcc	tactggatttt	tttatgctaa	cttccagttt	2700
tttagtattg	tttgatgcat	taaaaaggta	taggcctctg	ttatagttta	tgccaagtt	2760
atgagatgca	ttaatataca	ggggtccctg	ccccagttta	agacgtagtt	ttgtttgagc	2820
atcaaattggg	taatccacat	ctagaattaa	caagttgtta	tttatacgca	tgccaccgcc	2880
cgtttttaatt	tccatgttgt	ttgatgaatc	ataaccaata	gctcctgcaa	ctttgggtct	2940
aagggtgttt	tgttcaacgg	tgacacctgg	tccagtaact	actgttagtg	tatcggagtt	3000
ttgtgctact	tgcaaaggac	cgcttatttt	aatttctatt	tttccattat	ttacataaat	3060
aggatcttcc	atgttaatgc	ccaagctacc	cgtggcagta	gttagcgggg	gtgatgcagt	3120
tacagtaagg	gtgtcgctgt	cactgccaga	gaggggggct	gatgtttgca	gggctagctt	3180
tccatctgac	actgtaatgg	gccctttagt	agcaatgctt	agtttggagt	cttgcaagg	3240
cagtggggct	tgtgactgta	cgctaagagc	gccgctagta	actatcagag	gagcgggtgt	3300
tgccactggt	agggcgccctg	aggtaattgt	aagtgggtgcg	gaggtgtcca	aacttatggt	3360
tgactttgtt	tttttaagtg	gctgagtaac	agtgggtaca	ttttgggagg	tgaggtttcc	3420
ggccttgtct	agggtaagac	cgctgcccat	tttaagcgca	agcatgccgt	gggaggtgtc	3480
caaagggtcg	gagacgcgta	gagagagaac	tccaggggga	ctttcttgga	aaccattggg	3540
tgaacaaaat	ggaggggtaa	gaaagggcac	agttggaggc	ccggtttctg	tgcatatgg	3600
atacacgggg	ttgaagggtg	cttcagacgg	tctggcgcg	ttcatctgca	acaatatgaa	3660
gatagtgggt	cgggagggac	aagaacatga	ggaatttgac	atcccattta	aactttggag	3720
aaagtttgca	gctaaaaggc	ggctgagata	ccagagttgg	gaggaaggaa	aggaggtgat	3780
gctgaataag	ctggacaaag	atttgctgac	tgatttttaag	taagtaattt	attcagtcgt	3840
agccgtccgc	cgagtctttc	accgcgtcaa	agttgggaat	aaactgggtcc	gggtagtggc	3900
cgggaggtcc	agaaaagggg	ttgaagtaaa	ccgaaggcac	gaactcctca	ataaattgta	3960
gagttccaat	gcctccggag	cgcggtccg	aggacgaggt	ctgcagagtt	aggatcgct	4020
gacggggcgt	aaatgaagag	cggccagcgc	cgccgatctg	aaatgtccc	tccggacgga	4080

gaccaagaga	ggagctcacc	gactcgctcg	tgagctgaat	acctcgccct	ctgattttca	4140
ggtgagttat	accctgcccg	ggcgaccgca	ccctgtgacg	aaagccgccc	gcaagctgcg	4200
cccctgagtt	agtcactctga	acttcggcct	gggctgtctt	gggaagtacc	acagtgggtg	4260
gagcgggact	ttcctgttac	accagggcag	cgggccaact	acggggatta	aggttattac	4320
gaggtgtggt	ggtaatagcc	gcctgttcga	ggagaattcg	gtttcgggtg	gcgcggattc	4380
cgttgaccog	ggatatcatg	tggggctccg	cgctcatgta	gtttattcgg	gttgagtagt	4440
cttgggcagc	tccagccgca	agtcgccatt	gtggctggta	actccacatg	tagggcggtg	4500
gaatttcctt	gctcataatg	gcgctgacga	caggtgctgg	cgccgggtgt	ggccgctgga	4560
gatgacgtag	ttttcgcgct	taaattttgag	aaagggcgcg	aaactagctc	ttaagagtca	4620
gcgcgcagta	tttgctgaag	agagcctccg	cgtcttccag	cgtgcgcgca	agctgatctt	4680
cgcttttgtg	atacaggcag	ctgcgggtga	gggagcgag	agacctgttt	tttattttca	4740
gctcttggtc	ttggccctcg	ctttgttgaa	atatagcata	cagagtggga	aaaatcctat	4800
ttctaagctc	gcgggtcgat	acgggttcgt	tgggcgccag	acgcagcgct	cctcctcctg	4860
ctgctgccgc	cgctgtggat	ttcttgggct	ttgtcagagt	cttgctatcc	ggtcgccctt	4920
gcttctgtgt	gaccgctgct	gttgctgccc	ctgccgctgc	cgccgggtgca	gtaggggctg	4980
tagagatgac	ggtagtaatg	caggatgtta	cgggggaagg	ccacgcctgt	atggtagaga	5040
agaaagcggc	gggcgaagga	gatgttgccc	ccacagctct	gcaagcaagc	aactatggcg	5100
ttcttgtgcc	cgccgccacga	gcggtagcct	tggcgctgtt	gttgctcttg	ggctaaccggc	5160
ggcggctgct	tagacttacc	ggccctggtt	ccagtgggtg	cccactctacg	gttgggtcgg	5220
cgaacaggca	gtgcccggcg	cgcctgagga	gcggagggtt	tagcgatgct	gggaacgggt	5280
gccaatttct	ggggcgccgg	cgaggggaat	gcgaccgagg	gtgacgggtg	ttcgtctgac	5340
acctcttcgg	ctcgggaagc	ttcgtctagg	ctgtccagct	cttccatcat	ctcctcctcc	5400
tcgtccaaaa	cctcctctgc	ctgactgtcc	cagtattcct	cctcgtccgt	gggtggcggc	5460
ggcggcagct	gcagcttctt	tttgggtgcc	atcctgggaa	gcaagggccc	gcggctgctg	5520
atagggtctg	ggcggcgggg	ggattgggtt	gagctcctcg	ccggactggg	ggtccaggta	5580
aacccccctg	ccctttcgta	gcagaaactc	ttggcgggct	ttgttgatgg	cttgcaattg	5640
gccaaagatg	tggccctggg	taatgacgca	ggcggttaagc	tccgcatctg	gcgggcggga	5700
ttggtcttcg	tagaacctaa	tctcgtgggc	gtggtagtcc	tcaggtacaa	atttgccaag	5760
gtaagccgac	gtccacagcc	ccggagttag	tttcaacccc	ggagccgcgg	acttttcgtc	5820
aggcgaggga	ccctgcagct	caaaggtacc	gataatttga	ctttcgctaa	gcagttgcga	5880
attgcagacc	agggagcggt	gcgggggtgca	taggttgtag	cgacagttag	actccagtag	5940
gccgtcaccg	ctcacgtctt	ccatgatgtc	ggagtggtag	gcaaggtagt	tggctagctg	6000
cagaaggtag	cagtgaaccc	aaagcggcgg	agggcattca	cggtacttaa	tgggcacaaa	6060
gtcgctagga	agcgacagc	aggtggcggg	cagaattcct	gaacgctcta	ggataaagtt	6120
cctaaagttt	tgcaacatgc	tttgactggg	gaagtctggc	agaccctgtt	gcagggtttt	6180
aagcaggcgt	tcggggaaga	taatgtccgc	caggtgcgcg	gccacggagc	gctcgttgaa	6240
ggcgtcccat	aggtccttca	agttttgctt	tagcagcttc	tgcagctcct	ttaggttgcg	6300
ctcctccagg	cattgctgcc	acacgcccct	ggcgttttgc	caggtgtagc	acagaaataa	6360
gtaaacgcag	tcgcggacgt	agtcgcggcg	cgcctcgccc	ttgagcgtgg	aatgaagcac	6420
gttttgcccg	aggcggtttt	cgtgcaaaat	tccaaggtag	gagaccaggt	tgcagagctc	6480
cacgttgga	attttgcagg	cctggcgcac	gtagccctgg	cgaaggtgtg	agtgaacgt	6540
ttcctctagc	ttgcgctgca	tctccgggtc	cgctgcatgc	cgctgcatgc	actcaagctc	6600
cacggtaaca	agcactgcgg	ccatcattag	cttgctgcgc	tcctccaagt	cggcaggctc	6660
gcgcgtctca	agccagcgcg	ccagctgtct	atcgccaact	gcgggtaggc	cctcctcggt	6720
ttgttcttgc	aagtttgcac	ccctctccag	gggtcgtgca	cggcgcacga	tcagctcgct	6780
catgactgtg	ctcataacct	tggggggtag	gttaagtgcc	gggtaggcaa	agtgggtgac	6840
ctcgtatgtg	cgtttcagca	cggtatggcg	cggtttgtca	ccctcaagtt	ccaccagcac	6900
tccacagtga	ctttcatttt	cgctgttttc	ttgttgacga	gcgtttgccc	cgcgtttctc	6960
gtcgcgtcca	agaccctcaa	agatttttgg	cacttcgtcg	agcgaggcga	tatcagggtat	7020
gacagcgccc	tgccgcaagg	ccagctgctt	gtccgctcgg	ctgcgggttg	cacggcagga	7080
taggggtatc	ttgcagtttt	ggaaaaagat	gtgatagggt	gcaagcacct	ctggcacggc	7140
aaatacgggg	tagaagttga	ggcgcggtt	gggctcgcat	gtgccgtttt	cttggcggtt	7200
gggggtacg	cgcggtgaga	acaggtggcg	ttcgtaggca	aggctgacat	ccgctatggc	7260
gaggggcaca	tcgctgcgct	cttgcaacgc	gtcgcagata	atggcgcact	ggcgctgcag	7320
atgcttcaac	agcacgtcgt	ctccacatc	taggtagtcg	ccatgccttt	ggtccccccg	7380
cccgaactgt	tcctcgtttg	cctctgcgtc	gtcctggctc	tgetttttat	cctctgttgg	7440
tactgagcga	tcctcgctcg	cttcgcttac	aaaacctggg	tcctgctcga	taatcacttc	7500
ctcctcctca	agcgggggtg	cctcgacggg	gaaggtggta	ggcgcggttg	cggcatcggt	7560

ggaggcgggtg	gtggcgaaact	caaagggggc	ggttaggctg	tcctccttct	cgactgactc	7620
catgatcttt	ttctgcctat	aggagaagga	aatggccagt	cggaagagg	agcagcgcg	7680
aaccaccccc	gagcgcggac	gcggtgcggc	gcgacgtcca	ccaaccatgg	aggacgtgtc	7740
gtccccgtcg	ccgtcgccgc	cgcctccccg	cgcgccccca	aaaaagcggc	tgaggcggcg	7800
tctcagatcc	gaggacgaag	aagactcgtc	acaagatgcg	ctggtgcccgc	gcacaccag	7860
cccgcggcca	tcgacctga	cggcggattt	ggccattgcg	tccaaaaaga	aaaagaagcg	7920
cccctctccc	aagcccggag	gcccggcatc	ccagagggtg	atcgtggaca	gcgaggaaga	7980
aagagaagat	gtggcgctac	aaatggtggg	tttcagcaac	ccaccgggtg	taatcaagca	8040
cggcaagggg	ggtaagcgca	cgggtgcggcg	gctgaatgaa	gacgaccag	tggcgcgggg	8100
tatgcggacg	caagaggaaa	aggaagagtc	cagtgaagcg	gaaagtgaag	gcacgggtgat	8160
aaaccgcgtg	agcctgccga	tcgtgtctgc	gtgggagaag	ggcatggagg	ctgcgcgcgc	8220
gttgatggac	aagtaccacg	tggataacga	tctaaaggca	aacttcaagc	tactgcctga	8280
ccaagtggaa	gctctggcgg	ccgtatgcaa	gacctggcta	aacgaggagc	accgcgggtt	8340
gcagctgacc	ttcaccagca	acaagacctt	tgtgacgatg	atggggcgat	tcctgcaggc	8400
gtacctgcag	tcgtttgcag	aggtaacctt	caagcaccac	gagcccacgg	gctgcgcgtt	8460
gtggctgcac	cgctgcgctg	agatcgaaag	cgagcttaag	tgtctacacg	ggagcattat	8520
gataaataag	gagcacgtga	ttgaaatgga	tgtgacgagc	gaaaacgggc	agcgcgcgct	8580
gaaggagcag	tctagcaagg	ccaagatcgt	gaagaaccgg	tggggccgaa	atgtggtgca	8640
gatctccaac	accgacgcaa	ggtgctgctg	gcatgacgcg	gcctgtccgg	ccaatcagtt	8700
ttccggcaag	tcttgccgca	tggtcttctc	tgaaggcgca	aaggctcagg	tygctttttaa	8760
gcagatcaag	gctttcatgc	aggcgctgta	tcctaacgcc	cagaccgggc	acggtcacct	8820
tctgatgcca	ctacggtgcg	agtgcactc	aaagcctggg	catgcaccct	ttttgggaag	8880
gcagctacca	aagttgactc	cgttcgccct	gagcaacgcg	gaggacctgg	acgcggatct	8940
gatctccgac	aagagcgtgc	tggccagcgt	gcaccaccgc	gcctgatag	tgttccagtg	9000
ctgcaaccct	gtgtatcgca	actcgcgcg	gcagggcgga	ggccccaaact	gcgacttcaa	9060
gatatcgggc	cccgaacctg	taaacgcgtt	ggtgatgggt	cgcagcctgt	ggagtgaagaa	9120
cttcaccgag	ctgcgcggga	tggttgtgcc	tgagtttaag	tggagcacta	aacaccagta	9180
tcgcaacgtg	tcctgccag	tggcgcatag	cgatgcgcgg	cagaaccctt	ttgattttta	9240
aacggcgag	acggcaaggg	tggggggtaa	ataatcacc	gagagtgtac	aaataaaaaac	9300
atgtgccttt	attgaaagtg	tctcctagta	cattattttt	acatgttttt	caagtgaaca	9360
aaagaagtgg	cgctccta	ctgcgcactg	tggctgcgga	agtagggcga	gtggcgctcc	9420
aggaagtgtg	agagctgttc	ctggttgcga	cgcaggggtg	gctgtacctg	gggactgtta	9480
agcatggagt	tgggtacccc	ggtaataagg	ttcatgggtg	ggttgtgatc	catgggagtt	9540
tggggccagt	tggcaaaggc	gtggagaaac	atgcagcaga	atagtcacac	ggcggccgag	9600
ttgggcccct	gcacgccttg	ggtggacttt	tccagcggtt	tacagcggtc	gggggaagaa	9660
gcaatggcgc	tacggcgag	gagtgactcg	tactcaaact	ggtaaacctg	cttgagtcgt	9720
tggtcagaaa	agccaaagg	ctcaaagg	tagcatgttt	ttgagcgcg	gttccaggca	9780
aaggccatcc	agtgtagcc	cccagttctg	cgaccggccg	tattgactat	ggcgaggcg	9840
agcttggtg	gagaaacaaa	gcctggaaag	cgcttgatc	aggtgcccac	aaaatatggc	9900
ccacaacca	gatctttgac	aatggctttc	agttcctgct	cactggagcc	catggcggca	9960
gctgtgtgtg	atgtttgctg	cttcttttat	gttggtggcg	tgccggccga	gaagggcggtg	10020
cgcaggtaca	cggtctcgat	gacgcgcgg	tgcggctggt	gcacacggac	cacgtcaaa	10080
acttcaaa	aaacataaag	aagggtgggc	tcgtccatgg	gatccacctc	aaaagtcag	10140
tctagcgcgt	gggaggagtt	ggcgtagaga	aggttttggc	ccaggctctg	gagtgccccc	10200
atggacataa	agttactgga	gaatgggatg	cgccaaagg	tgcgatcgca	aagaaacttt	10260
ttctgggtaa	tactgtcaac	cgcggttttg	cctattagtg	ggtagggcac	gttggcgggg	10320
taagcctgtc	cctcgcgcat	ggtgggagcg	aggtagccta	cgaatcctga	gttggttatg	10380
tggtaagaa	ttccaacctg	ctgatactcc	ttgtatttag	tatcgtcaac	cacttgccgg	10440
ctcatgggct	ggaagtttct	gaagaacgag	tacatgcggt	ccttgtagct	ttctggaatg	10500
tagaagccct	ggtagccaat	attgtagttg	gccaacatct	gcaccaggaa	ccagtccttg	10560
gtcatgttgc	actgagctac	gttgtagccc	tccccgtcaa	ctgagcggtt	aatctcaaac	10620
tcattgggag	taagcaggcg	gtcgttgccc	ggccagctaa	cagaagagtc	aaaggtaatg	10680
gccaccttct	taaagggtg	attaagatag	aagggtccgt	caaggatatg	tatggagcca	10740
gagtaggtgt	agtaagggtc	gtagcctgat	cccagggaag	gggtttcctt	tgtcttcaag	10800
cgtgtgaagg	cccaaccgcg	aaatgctgcc	cagttgcgcg	atgggatgga	gatggggcac	10860
ttggtggcgt	tggcgggtat	ggggtatagc	atgttgccgg	cggaaaggta	gtcattaaag	10920
gactggtcgt	tgggtgtcatt	tctgagcatg	gcttcacgcg	tggaggccgt	gttggtgggc	10980
atggggaaga	aggtggcgta	aagacaaatg	ctgtcaaac	taatgctagc	cccgtaact	11040

ctaagatcgt	ttcccagaga	gctctgcaga	accatgttaa	catccttcct	gaagttccat	11100
tcatatgtat	atgagcctgg	caggaggagg	agggttttaa	tggcaaaaaa	cttttggggc	11160
acctgaatgt	gaaagggcac	gtagcggccg	tttcccaaca	acatggagcg	ataacggagg	11220
cccgcatgtc	ggtggtggtt	aaagggatta	acgttgtcca	tgtagtccag	agaccagcgc	11280
gccccaaagt	taatgtagca	gtctacaagc	ccgggagcca	ccactcgctt	gttcatgtag	11340
tcgtagggtg	tggggttgtc	agatatattc	acattggtgg	ggttgatatt	tagcttgtct	11400
ggcagggtaca	gcgcaatatt	ggagtaaagg	aaattttctc	atagggtggc	atttaggtta	11460
atttccatgg	caaagtgtgt	accacttcct	atttcattac	gtgttgcaaa	agtttcatct	11520
tttgtccatg	tagtatctcc	attatcgcc	gagccattgc	cattagcctt	aatagcttga	11580
taggtgtcag	ttaccccaat	acccccaaga	ggaaaacaat	aatttggcaa	ttcatcctca	11640
gttccatggg	tttcaatgat	tctaacatct	ggatcatagc	tgtctacagc	ctgattccac	11700
atagaaaaat	atctggttct	atcacctatg	gaatcaagca	agagttgata	ggacagctct	11760
gtgtttctgt	cttgcaaatc	taccacggca	tttagctgcg	atgcctgacc	agcaagaaca	11820
cccatgttgc	cagtgcgtgt	ataatacatt	aggccaataa	aattgtccct	gaaagcaatg	11880
taattgggtc	tgtttggcat	agattgttga	cccaacatag	ctttagaatt	ttcatcacct	11940
tttccagggt	tgtaaagacag	atgtgtgtct	ggggtttcca	tatttacatc	ttcactgtac	12000
aaaaccactt	ttggtttagt	agcattgcct	tgcgggtcgt	tcaaagaggt	agtatttgag	12060
aagaattgca	agtcaacctt	tgggaagagg	accccttttt	catccggaac	cagaacggat	12120
tgaccaccaa	aaggatttgt	aggcctggca	taagatccat	agcatgggtt	catggggagt	12180
gtttttttaa	gcactctccc	tcctgcccga	ttagcatcag	cttcggttcca	ctgagattcg	12240
ccaatttgag	gttctggttg	ataggaagga	ctgcgtata	cagggttagc	ttgtgtttct	12300
gcattgtctg	atcctatttg	tagcccgtt	tttgtaattg	tttctccaga	caaaggagcc	12360
tgggcataga	catgtgtttt	cttagtagcc	tgatctcgag	cgttttgctc	ttcttcttcc	12420
tcttcttcat	cttcatcttc	ctcttcttca	tcctcggcaa	ctgcccggcc	gctatcttcg	12480
gtttgttccc	actcacagga	gttaggagcg	cccttgggag	ctagagcggt	gtaggcagtg	12540
ccggagtagg	gcttaaaagt	aggccctctg	tcagcacgc	cgcggatgtc	aaagtacgtg	12600
gaagcatat	caagcacacg	gttgtcaccc	acagccaggg	tgaaccgcgc	tttgtacgag	12660
tacgcggtat	cctcgcggtc	cacagggatg	aaccgcagcg	tcaaacgctg	ggaccggctc	12720
gtggttacgt	cgtgcgtagg	tgccaccgtg	gggtttctaa	acttgttatt	caggctgaag	12780
tacgtctcgg	tggcgcgggc	aaactgcacc	agcccggggc	tcagggtactc	cgaggcgctc	12840
tggcccagaga	tgtgcatgta	agccactgc	ggcatcatcg	aaggggtagc	catcttgtaa	12900
agcgggcgca	cggcggtca	gcagctcctc	tggcggcgac	atggacgcat	acatgacaca	12960
tacgacacgt	tagctattta	gaagcatcgt	cggcgcttca	gggattgcac	ccccagaccc	13020
acgatgctgt	tcagtgtgct	ttgccagttg	ccactggcta	cgggcgcgat	cgatcgcgga	13080
ccgctggcgg	cacggcgagc	ggacgcgcgg	ctagggcggg	ttacaacaac	ggcggacggc	13140
cctggcagca	caggtttctg	ctgggtgtca	gcggggggag	gcagggtccag	cgttacaggt	13200
gtgtgctggc	ccagcactcc	ggtagccatg	ggcgcgatgg	gacgggtggt	gggcaggcct	13260
tgtcttagtg	cctcctcgta	cgagggaggg	tcacttattt	gcgtcaccag	agtttcttcc	13320
ctgtcggggc	gcgacgcctt	ttcgccacgc	ccctctggag	acactgtctc	cacggccggg	13380
ggaggctcct	ctacgggagg	gcggggatca	agcttactgt	taatcttatt	ttgactggc	13440
tggttggcca	ggtccaccac	cccgtctaat	ccagaggcca	ggccatctac	cactttttgt	13500
tggaaatttt	gctctttcaa	cttgtccctc	agcatctggc	ctgtgctgct	gttccaggcc	13560
ttgctgccat	agttcttaat	ggtggaaccg	aaatttttaa	tgcgcgtcca	cagcgagccc	13620
cagctgaagg	cgccaccgct	catattgctg	gtgccgatat	cttgccagtt	tcccatgaac	13680
gggcgcgagc	cgtgtcgcgg	ggccagagac	gcaaagttga	tgtcttccat	tctacaaaat	13740
agttacagga	ccaagcgagc	gtgagactcc	agacttttta	ttttgatttt	tccacatgca	13800
acttgttttt	aatcagtgtc	tctgcgcctg	caaggccacg	gatgcaattc	cgggcacggc	13860
gccaatcgcc	gcggcgatca	gtggaataag	gaggggaggg	ataccgcccgc	gcatgcgacg	13920
gtgcgacgcg	cgccgcgcgc	ggtggtgcgc	acgacgcgat	ccgcccgtca	ggccgtggcc	13980
ggccatgccc	ctcctacggg	gcattcttcc	tccgaatccc	ggcaccggga	aacggaggcg	14040
gcaggtgagg	gccatatctg	caagaaccac	aaagaccggc	ttttaaacga	tgctgggggtg	14100
gtagcgcgct	gttggcagca	ccagggtcct	gcctccttcg	cgagccaccc	tgcgacagga	14160
aatcggggcc	agcacgggct	ggcgacggcg	acggcgggcg	cgggttccag	tgggtggttcg	14220
gcgtcgggta	gtcgtctcgtc	ttctggggcg	gtaggtgtag	ccacgatagc	cgggggtagg	14280
cgcatgggaa	ggatgtaggg	catattcggg	cagtagtgcg	ctggcggtgc	cgtacttctc	14340
ggaacggcgc	gggcgcgggg	gggctgaaac	gcgaaacatc	cacgggtccg	tttgacctc	14400
cgtagagggt	ttggacggcg	ccgcagcgcg	cgcctgcacc	gcggcatctg	ccaccgcca	14460
ggcaaccggg	gacgtttgtg	tctccatgcc	ctctgtggca	gtggcaatac	tagtgctact	14520

ggtggtgggt	atctgaacgt	ccacgggtctg	cacgcccagt	cccgggtgcca	cctgcttgat	14580
tggccgcacg	cggacctcgg	gctccagccc	aggctccacg	gtcatttttt	ccaagacatc	14640
ttccagtcgc	tggcgcttgg	gtaccatcag	ctgcacgggtg	ggtgccaaagt	caccagactc	14700
gcgctttagg	ccgcgccttt	cttcggacgg	tgcaagcggtg	ggcagcacct	gctgcagtgt	14760
cacgggcttt	aggctaggtg	ttgggttgcc	ctcgtccagc	ggcaacgcca	acatgtcctt	14820
atgccgcttt	ccgtaggcaa	actccccgag	gcgctcgttg	gcctgctcaa	gcaggctcct	14880
gtcgcggtac	acctcatcat	acacgcgctt	gtaggtgctg	gtggagcgct	caccgggctg	14940
aaaaactacg	gtggtgccgg	gtcgcaaaac	acgtcttacg	cgtagacctt	tccactgtac	15000
ccgcgcgctg	ggcgcggttg	cgtgcagcag	ttccacctcg	tcgtcaagtt	catcatcatc	15060
atcatctttc	tttttctttt	tgacctgctt	tagctttcgg	ggcttgtaat	cctgctcttc	15120
cttcttcggg	gggcataga	tctccggcgc	gatgacctgg	agcatctctt	ctttgatttt	15180
gcgcttgga	atagcttctg	tgcgcgcgcg	cgccgctgga	tacatacaac	agtacgagtc	15240
taagtagttt	tttcttgcaa	tctagttgctg	cggggggcgg	gtgcgcacgg	gcacgcgcag	15300
gccgctaacc	gagtcgcgca	cccagtacac	gttgccctcg	cgacctgag	tcatagcaat	15360
aatggccgcg	gctgctgcgg	cggccgctcg	tcgcctggac	ctggggggca	cagtgaaca	15420
accgcggg	agccttcgag	cggcccgcat	ggccgcctgt	cggccgggtg	gacgtgcgcg	15480
gctaagcagg	gcgcgcgg	cgcttgggc	ggcagtgccg	ggtcgggcgg	ggtggcgagc	15540
tgctacgcgc	ctccgcgcgc	tcttcatttt	agcataacgc	cgggctccgc	gcaccacggg	15600
ctgaatggcc	gcgtccactg	tggacactgg	tggcgcgctg	ggcgtgtagt	tgcgcgccct	15660
ctccaccacc	gcgtcaatgg	cgtcatcgac	ggtggtgctg	ccagtgcggc	cgcttctgtg	15720
cgcgccccag	ggcgcgcggg	agtgcgcgcg	cacgcgcact	gggtgttggt	cggagcgctt	15780
ctttgccccg	ccaaacatct	tgcttgggaa	cgcaggccc	cagcctgtgt	tattgctggg	15840
cgatataagg	atggacatgt	ttgctcaaaa	agtgcggctc	gataggacgc	gcggcgagac	15900
tatgcccagg	gccttgtaaa	cgtaggggca	ggtgcggcgt	ctggcgctcag	taatggtcac	15960
tcgctggact	cctccgatgc	tggtgcgcag	cggtagcgct	ccgtgatctg	tgagagcagg	16020
aacgttttca	ctgacggtgg	tgatggtggg	ggctggcggg	cgcgccaaaa	tctggttctc	16080
gggaaagcga	ttgaacacgt	gggtcagaga	ggtaaactgg	cggatgagct	gggagtagac	16140
ggcctggtcg	ttgtagaagc	tcttgaggtg	cacgggcaac	agctcggcgc	ccaccaccgg	16200
aaagtgtgtg	atctggctcg	tggagcgga	ggtcacgggg	tcttgcatca	tgtctggcaa	16260
cgaccagtag	acctgctccg	agccgcagg	tacgtcagga	gtgcaaagga	gggtccatga	16320
gcggatcccg	gtctgaggg	cgccgtagtt	gtatgcaagg	taccagctgc	ggtactgggt	16380
gaaggtgctg	tcattgctta	ttaggttgta	actgcgtttc	ttgctgtcct	ctgtcagggg	16440
tttgatcacc	ggtttcttct	gaggttcttc	gacctcgggt	tgccagcgcg	gggcggcagc	16500
ttctgcccgt	gcctcggcct	cagcgcgctt	ctcctccgcc	cgtgtggcaa	aggtgtcgcc	16560
gcgaatggca	tgatcgttca	tgctctccac	cggctgcatt	gccgcggctg	ccgcgttgga	16620
gttctcttcc	gogccgctgc	cactgttggt	gccgcgcctt	gcgccatccc	cgcctgttcc	16680
ggtgtcatct	tttaagcttg	cctggtaggc	gtccacatcc	aacagtgcgg	gaatggtacc	16740
acctccagg	tcacgttagg	tgatcctaaa	gccctcctgg	aagggttgcc	gcttgcggtg	16800
gcccacaag	ttgctcaggc	ggctgtgggt	gaagtccacc	ccgcatcctg	gcagcaaat	16860
gatgtctgga	tgggaaggctt	cgtttgtata	taccccaggc	atgacaagac	cagtgaactg	16920
gtcaaacccc	agtctgaagt	tgcgggtgtc	aaactttacc	ccgatgtcgc	tttccagaac	16980
cccgttctgc	ctgcccactt	tcaagtagtg	ctccacgatc	gcgttggtca	taagggtctat	17040
ggtcatgggt	tcggagtagt	tgccctcggg	cagcgtgaac	tccaccact	catatttcag	17100
ctccacctgt	ttgtccttag	taagcgagcg	cgacaccatc	accgcgcct	taaaacttatt	17160
ggtaaacatg	aactcgttca	catttggtcat	gttggtatgc	aggatgggtt	tcagggtcgcc	17220
gccccagtg	gaacggtcgt	caagattgat	ggtctgtgtg	cttgccctcc	ccgggctgta	17280
gtcattgttt	tgaatgaccg	tggttagaaa	gttgctgtgg	tcgttctggg	agttcagggg	17340
tgccacatcc	ggtgacttgt	tgtccacaag	gtacacacgg	gtggtgtcga	ataggggtgc	17400
caactcagag	taacggatgc	tgtttctccc	cccggtaggc	cgcagggtacc	gcggaggcac	17460
aaacggcggg	tccaggggag	catcgaagg	ggaacccagc	gccgcgcgca	ctggcgccgc	17520
gctcaccacg	ctctcgtagg	agggaggagg	accttccctca	tacatcgccg	cgcgctgcat	17580
actaagggga	atacaagaaa	accaacgctc	ggtgccatgg	ccttggtgag	ttttttattt	17640
tgcatcatgc	tttttttttt	tttttttaaa	acatttctcc	cagcctgggg	cgaagggtgcg	17700
caaacgggtt	gccactccct	cccaaataca	ggacgctgct	gtcgtctgcc	gagtcacgt	17760
cctcccacac	cagaccccg	tgacggtcgt	gcctttgacg	acgggtgggc	gggcgcgggc	17820
cgggcacatc	cctgtgctcc	tgcgcatacg	tcttccatct	actcatcttg	tccactaggc	17880
tctctatccc	ggtgttgga	aatgcgggag	gcaggttctt	ttcgcgctgc	ggctgcagca	17940
gcgagttgtt	taggtactcc	tcctcgccca	gcaggcgcg	gcgggtgggtg	cgagtgcctg	18000

taaaagaccc	tatcaagctt	ggaaatgggc	tactcgcatc	tgaccgcggg	gccgcagcgc	18060
ctagatcgga	caagctgctt	ggcctgcgga	agctttcctt	tcgcagcgcc	gcctctgcct	18120
gctcgcgctg	ttgcaactct	agcagggctc	gcggttgctg	ggaaaacacg	ctgtcgctcta	18180
tgctgtccca	gaggaatcca	tcgttaccct	cgggcacctc	aaatcccccg	gtgtagaaac	18240
cagggggcgg	tagccagtgc	gggttcaaga	tggcattggt	gaaatactcg	gggttcacgg	18300
cggccgcgcg	atgcaagtag	tccattagcg	gattgataaa	cggccggttt	gaggcataca	18360
tgcccggttc	catgttgctg	gcgggtcatgt	ccagcgccac	gctgggctgt	accccgctgc	18420
gcatcaggtt	aaggctcacg	ctctgctgca	catagcgcaa	gatgcgctcc	tcctcgctgt	18480
ttaaactgtg	caacgagggg	atcttctgcc	gccggttggt	cagcaggtag	ttcaggggtg	18540
cctccagget	gcccgtgtcc	tcctgcccc	gcgcgcggct	gacacttgta	atctcctgga	18600
aagtatgtc	gtccacatgc	gcctgacct	tggcctcgcg	gtacagtgtc	agcaagtgc	18660
ctaggtatgt	gtcccggtac	acgtgccac	tgtccgtgaa	gggcgtatt	agcagcagca	18720
acaggcgcg	gttgggctgc	agcaagctag	acacggctgc	gcggctgcct	gtgggagccc	18780
gcacccccca	cagccccctg	aagttcttga	aagcctggct	caggtttacg	gtctgcaggc	18840
cttgtctact	ggtctggaaa	aaatagtctg	gcccggactg	gtacacctca	ctttgcggtg	18900
tctcagtcac	cattagccgc	agtgcgtca	caaagttggt	gtagtctctc	tgtccccgcg	18960
gcacgttggc	ggctgtgtga	ctcaggaagg	cgtttagtgc	aacctggag	cccaggttgc	19020
cctgtctgtg	cgcgcgctca	cgctgcgcca	cggcctcgcg	cacatcccc	accagccggt	19080
ccaggttggg	ctgcacgttg	ccgtgttgt	aacgagccac	gcgtgaagc	agcgcgctgt	19140
agaccaggcc	ggcctcatcg	ggcggatgg	ccctgttttc	ggccagcgcg	tttacgatcg	19200
ccagcacctt	ctcgtgcgtg	gggtttgcgc	gcgcgggac	caccgcttcc	agaattgcga	19260
agagccgggt	ggcctgcgca	tgctgcggga	acgcgtcagg	gttacgcgca	gtcagcgaca	19320
tgatgcggtc	catgacctgg	cgccagtcgt	ccgtggagtt	aaggccggac	ggctggctct	19380
gcagcgccgc	cgcaccgcgc	gggtccgttg	cgtcttgcat	catctgatca	gaaacatcac	19440
cgcttagtac	tcgccgtcct	ctggctcgta	ctcatcgctc	tcgtcatatt	cctccacgcc	19500
gccgacgttg	ccagcgcgcg	cgggtgccac	cgccagccca	gggtccggccc	cagctgcctc	19560
cagggcgcgt	cgcttgggg	cccagcgag	gtcagcgccc	gcgtcaaagt	aggactcggc	19620
ctctctatcg	ccgtgcccgc	tgccagccag	ggccctttgc	aggetgtgca	tcagctcgcg	19680
gtcgctgagc	tcgcgcgcgc	ggctcacgct	cacggccttg	tggatgcgct	cgttgcgata	19740
aacgcccagg	tcgtcgctca	aggtaagcac	cttcaacgcc	atgcgcattg	agaaccctc	19800
gatctttacc	tccttgtcta	tgggaacgta	aggggtatgg	tatatcttgc	gggcgtaaaa	19860
cttgcccaga	ctgagcatgg	aatagttaat	ggcgccacc	ttgtcagcca	ggctcaagct	19920
gcgctectgc	accactatgc	tctgcagaat	gtttatcaaa	tcgagcagcc	agcgccctc	19980
gggctctact	atgttttagca	gcgcacccct	gaatgcctcg	ttgtccctgc	tgtgctgcac	20040
tataaggaac	agctgcgcca	tgagcggctt	gctatttggg	ttttgtcca	gcgcgcttac	20100
aaagtccac	agatgcatca	gtcctatagc	cacctctcg	cgcgccacaa	gcgtgcgca	20160
gtggttggtt	aagctttttt	gaaagttaat	ctcctggttc	accgtctgct	cgtagcggt	20220
taccaggtcg	gcggccgcca	cgtgtgcgcg	cgcgggacta	atcccggtcc	gcgcgtcggt	20280
ctcaaagtcc	tcctcgcgca	gcaaccgctc	gcgggttcagg	ccatgcgcga	actcgcgccc	20340
tgcggtgaac	tttcgatccc	gcattctctc	gggtcctctc	ccctcgcggt	cgcgaaacag	20400
gttctgcccgc	ggcacgtacg	cctcgcgctg	gtcacgcttc	agctgcaccc	ttgggtgtcg	20460
ctcaggagag	ggcgctccta	gccgcgccag	gcctcgccc	tcctccaagt	ccaggtatgt	20520
ccgggcccgcg	cgcgcggggg	gttcgtaatc	accatctgcc	gccgcgtcag	ccgcggatgt	20580
tgccctcct	gacgcggtag	gagaagggga	gggtgcccctg	catgtctgcc	gctgctcttg	20640
ctcttgccgc	tgctgaggag	gggggcgcac	ctgccgcagc	accggatgca	tctgggaaaa	20700
gcaaaaaagg	ggctcgctcc	tgtttccgga	ggaatttgca	agcggggtct	tgcatgacgg	20760
ggaggcaaac	ccccgttcgc	cgcagtcgag	ccggcccag	actcgaaccg	ggggtcctgc	20820
gactcaaccc	ttggaaaata	accctccggc	tacagggagc	gagccactta	atgctttcgc	20880
tttcagacct	aaccgcttac	gccgcgcgcg	gccagtggcc	aaaaaagcta	gcgcagcagc	20940
cgcgcgcct	ggaaggaagc	caaaaggagc	gctccccctg	tgtctgacgt	cgcacacctg	21000
ggttcgacac	gcgggcggtg	accgcattga	tcacggcgga	cggccggatc	cggggttcga	21060
accccgctcg	tcgcccata	tacccttgcg	aatttatcca	ccagaccacg	gaagagtgc	21120
cgcttacagc	ctctcctttt	gcacggtcta	gagcgtcaac	gactgcgcac	gcctcaccgg	21180
ccagagcgtc	ccgacctagg	agcacttttt	gccgctgcgc	aacatctgga	accgcgtccg	21240
cgactttccg	cgcgcctcca	ccaccgcgcg	cggcattacc	tggatgtcca	ggtacatcta	21300
cggatatcat	cgccttatgt	tggaaagacct	cgcgcccgga	gccccggcca	ccctacgctg	21360
gccccctac	cgcagccgcg	cgcgcgactt	tttggtggga	tatcagtacc	tggtcgggac	21420
ttgcaacgac	tacgtctttg	actcaagggc	ttactcgctg	ctcaggtaca	ccgagctctc	21480

gcagccgggt	caccagaccg	ttacttggtc	cgttatggcc	aactgcactt	acaccatcaa	21540
cacgggcgca	taccaccgct	ttgtggacat	ggatgacttc	cagtctaccc	tcacgcaggt	21600
gcagcaggcc	atattagccg	agcgcgttgt	cgccgacctg	gccctgcttc	agccgatgag	21660
gggcttcggg	gtcacacgca	tgggaggaag	agggcgccac	ctacggccaa	actccgccgc	21720
cgccgtagcg	atagatgcaa	gagatgcagg	acaagaggaa	ggagaagaag	aagtgccggt	21780
agaaaggctc	atgcaagact	actacaaaga	cctgcgccga	tgtcaaaacg	aagcctgggg	21840
catggccgac	cgcctgcgca	ttcagcaggc	cggacccaag	gacatggtgc	ttctgtcgac	21900
catccgccgt	ctcaagaccg	cctactttaa	ttacatcatc	agcagcacct	ccgccagaaa	21960
caaccccgac	cgccaccgcg	tgcgcgccgc	cacggtgctc	agcctacctt	gcgactgtga	22020
ctggttagac	gcctttctcg	agagggtttc	cgatccggtc	gatgcggact	cgctcaggtc	22080
cctcggtggc	ggagtaccta	cacaacaatt	gtttagatgc	atcgttagcg	ccgtatccct	22140
gccgcacggc	agcccccg	caacccataa	ccgggacatg	acgggcggcg	tcttccaact	22200
gcgccccgcg	gagaacggcc	gcgcggtcac	cgagaccatg	cgccgtcgcc	gcggggagat	22260
gatcgagcgc	tttgtcgacc	gcctcccggt	gcgcggtcgt	cgccgccgtg	tccccctcc	22320
cccaccgccg	ccagaagaag	aagaagaagg	ggaggccctt	atggaagagg	agattgaaga	22380
agaagaggcc	cctgtagcct	ttgagcgca	ggtgcgcgac	actgtcgccg	agctcatccg	22440
tcttctggag	gaggagttaa	ccgtgtcgcc	gcgcaactcc	cagtttttca	acttcgccgt	22500
ggactttctac	gaggccatgg	agcgccttga	ggccttgggg	gatatcaacg	aatccacggt	22560
gcgacgtggg	gttatgtact	tcttcgtggc	agaacacacc	gccaccaccc	tcaactacct	22620
ctttcagcgc	ctgcgaaact	acgcgctctt	cgcccgycac	gtggagctca	atctcgcgca	22680
ggtggtcatg	cgcgcgccgc	atgccgaagg	gggctgtgtc	tacagccgcg	tctggaacga	22740
gggaggcctc	aacgccttct	cgcagctcat	cgcccgcatc	tccaacgacc	tcgccgccac	22800
cgtggagcga	gccggacgcg	gagatctcca	ggaggaagag	atcgagcagt	tcattggccga	22860
aatcgccctat	caagacaact	caggagacgt	gcaggagatt	ttgcgccagg	ccgccgtcaa	22920
cgacaccgaa	attgattctg	tcgaactctc	tttcagggtc	aagctcaccg	ggcccgctcg	22980
cttcacgcag	aggcgccaga	ttcaggagat	caaccgccgc	gtcgtcgctg	tcgccagcaa	23040
cctccgcgcg	cagcaccagc	tccctgccgc	gcgcggcgcc	gacgtgcccc	tgccccctct	23100
cccggccgggt	cccagacccc	ccctacctcc	gggggccccg	ccgcgtcacc	gcttttagat	23160
gcacatcca	aggacacccc	cgcggccccc	cgcccgccgc	gcggtaccgt	agtcgcgcgc	23220
cggggatgcg	gcctcttgca	agtcacgcac	gccgccacca	accagccccct	ggaaatcagg	23280
tatcacctgg	acctagcccc	cgccttgacc	cggctatgcg	aggtaaacct	gcaggagctc	23340
ccgctcgacc	tgtcgccgcg	ggagctccag	accatggaca	gctcccatct	gcgcgatggt	23400
gtcatcaagc	tccgaccgcc	gcgcgcggac	atctggactt	tgggctcgcg	cgccgtgggtg	23460
gtccgatcca	ccataactcc	cctcgagcag	ccagacggtc	aaggacaagc	agccgaagta	23520
gaagaccacc	agccaaaccc	gccaggcgag	gggtcacaat	tcccactctg	cttccctgtg	23580
cgcggtcgct	aggtcaacct	cgtgcaggat	gtacagcccg	tgcaccgctg	ccagtactgc	23640
gcacgttttt	acaaaagcca	gcacgagtgt	tcggcccgtc	gcagggaactt	ctactttcac	23700
cacatcaaca	gccactcctc	caactggtgg	cgggagatcc	agttcttccc	gatcggtctg	23760
catcctcgca	ccgagcgtct	ctttgtcacc	tacgatgtag	agacctatac	ttggatgggg	23820
gcctttggga	agcagctcgt	gcccttcacg	ctggttatga	agttcggcgg	agatgagcct	23880
ctggtgaccg	ccgcgcgaga	cctagccgtg	gaccttggtg	gggaccgctg	ggaacaagac	23940
ccgcttacct	tctactgcat	caccccagaa	aaaatggcca	taggtcgcca	gtttaggacc	24000
tttcgcgacc	acctgcaaat	gctaattggc	cgtgacctgt	ggagctcatt	cgtcgcttcc	24060
aacctctatc	ttgcagactg	ggcctgttca	gaacacgggc	tcagctcccc	tgaggagctc	24120
acctacgagg	aacttaaaaa	attgccctcc	atcaagggca	ccccgcgctt	cttggaaactt	24180
tacatcgtag	gccacaacat	caacggcttc	gacgagatcg	tgtcgcgcgc	ccaggtaatt	24240
aacaaccggt	ccgaggtgcc	gggacccttc	cgcacacac	gcaactttat	gcctcgcgcg	24300
ggaaagatac	ttttcaacga	tgtcaccttc	gccctgccaa	acccgcgctt	caaaaagcgc	24360
acggactttt	tgtcttgagg	gcaggggcga	tgcgacgaca	ctgacttcaa	ataccagtac	24420
ctcaaagtca	tggttaggga	cacctttcgc	ctcaccacac	cctcgctccg	gaaggccgcg	24480
caggcatacg	cgctaccggt	agaaaaggga	tgtcgcgcct	accaggccgt	caaccagttc	24540
tacatgctag	gctcttacgc	ttcggaggcc	gacgggtttc	cgatccaaga	gtactggaaa	24600
gaccgcgaag	agttttgtcct	caaccgcgag	ctgtggaaaa	aaaagggaaca	ggataagtat	24660
gacatcatca	aggaaaccct	ggactactgc	gccctagacg	tgcagggtcac	cgccgagctg	24720
gtcaacaagc	tgcgcgactc	ctacgcctcc	ttcgtgcgtg	acgcggtagg	tctcacagac	24780
gccagcttca	acgtcttcca	gcgtccaacc	atatcatcca	actcacatgc	catcttcagg	24840
cagatagtct	tccgagcaga	gcagcccgcc	cgtagcaacc	tcgggtccga	cctcctcgct	24900
ccctcgcaag	aactatacga	ttacgtgcgc	gccagcatcc	gcggtggaag	atgctaccct	24960

acatatcttg	gaatactcag	agagccctc	tacgtttacg	acatttgcg	catgtacgcc	25020
tccgcgctca	cccccccat	gccatgggg	ccccactca	accatacga	gcgcgcgctt	25080
gccgcccgcg	catggcagca	ggcgctagac	ttgcaaggat	gcaagataga	ctacttcgac	25140
gcgcgcctgc	tgcccgggg	ctttaccgtg	gacgcagacc	ccccggacga	gacgcagcta	25200
gacccactac	cgccattctg	ttcgcgcaag	ggcgcccgcc	tctgctggac	caacgagcgc	25260
ctacgcggag	aggtagccac	cagcgttgac	cttgtcacc	tgacaaccg	cggttggcgc	25320
gtgcacctgg	tgcccgacga	gcgcaccacc	gtctttcccg	aatggcggtg	cgttgcgcg	25380
gaatacgtgc	agctaaacat	cgcgcccaag	gagcgcgccg	atcgcgacaa	aaaccaaacc	25440
ctgcgcctcca	tcgccaagtt	gctgtccaac	gccctctacg	ggtcggttgc	caccaagctt	25500
gacaacaaaa	agattgtctt	ttctgaccag	atggacgcgg	ccaccctcaa	aggcatcacc	25560
gcgggccagg	tgaatatcaa	atcctcctcg	tttttggaaa	ctgacaatct	tagcgagaa	25620
gtcatgcccc	cttttgagag	ggagtactca	ccccaacagc	tggccctcgc	agacagcgat	25680
gcggaagaga	gtgaggacga	acgcgcccc	acccctttt	atagccccc	ttcaggaaca	25740
cccggtcacg	tgccctacac	ctataaacca	atcaccttc	ttgatgccga	agagggcgac	25800
atgtgtcttc	acaccctgga	gcgagtggac	cccctagtgg	acaacgaccg	ctaccctcc	25860
cacttagcct	ccttcgtgct	ggcctggacg	cgagccttcg	tctcagagt	gtccgagttt	25920
ctatagcgg	aggaccgcg	aacaccgctc	gaggacaggc	ctctcaagtc	tgtatcggg	25980
gacacggaca	gccttttctg	caccgagcgt	ggacaccggc	tcatggaaac	cagaggtaag	26040
aaacgcacat	aaaagcatgg	gggaaacctg	gtttttgacc	ccgaacggcc	agagctcacc	26100
tggtctgctg	aatgcgagac	cgctctcggg	gcctgcggcg	cggtatgcta	ctccccggaa	26160
tcggtatttc	tcgcgcccc	gctctacgcc	cttaaaagtc	tgactgccc	ctcgtgcggc	26220
gcctcctcca	aggcaagct	gcgcgccaag	ggccacggcg	cggaggggct	ggactatgac	26280
accatggtca	aatgctacct	ggccgacgcg	cagggcggaag	accggcagcg	cttcagcacc	26340
agcaggacca	gcctcaagcg	caccctggcc	agcgcgagc	ccggagcgca	ccccttcacc	26400
gtgacccaga	ctacgctgac	gaggaccctg	cgcccgtgga	aagacatgac	cctggcccgt	26460
ctggacgagc	accgactact	gccgtacagc	gaaagccgcc	ccaaccgcg	aaacgaggag	26520
atatgctgga	tcgagatgcc	gtagagcacg	tgaccgagct	gtgggaccgc	ctggaactgc	26580
ttggtcaaac	gctcaaaagc	atgcctacgg	cggacggcct	caaaccgttg	aaaaactttg	26640
cttccttgca	agaactgcta	tcgctggcg	gcgagcgct	tctggcgcat	ttggtcagg	26700
aaaacatgca	agtcaaggac	atgcttaacg	aagtggcccc	cctgctcagg	gatgacggca	26760
gctgcagctc	tcttaactac	cagttgcagc	cggtaatagg	tgtgatttac	gggcccaccg	26820
gctgcggtaa	gtcgcagctg	ctcaggaacc	tgctttcttc	ccagctgac	tcccctacc	26880
cggaaacgg	tttcttcac	gccccgcagg	tagacatgat	cccccatct	gaactcaaag	26940
cgtgggaaat	gcaaatctgt	gagggtaact	acgcccctgg	gccggatgga	accattatac	27000
cgcagctctg	caccctccgc	ccgcgctttg	taaaaatggc	ctatgacgat	ctcatcctgg	27060
aacacaacta	tgacgttagt	gatcccagaa	atatcttcgc	ccaggccgcc	gcccgtgggc	27120
ccattgccat	cattatggac	gaatgcatgg	aaaatctcgg	aggtcacaa	ggcgttcca	27180
agttcttcca	cgcatctcct	tctaagctac	atgacaaatt	tcccaagtgc	accggatata	27240
ctgtgctgg	ggttctgcac	aacatgaatc	cccggaggga	tatggctggg	aacatagcca	27300
acctaaaaat	acagtccaag	atgcatctca	tatccccacg	tatgcacca	tcccagctta	27360
accgctttgt	aaacacttac	accaagggcc	tgcccctggc	aatcagcttg	ctactgaaag	27420
acatttttag	gcaccacgcc	cagcgctcct	ctacgactg	gatcatctac	aacaccacc	27480
cgcagcatga	agctctgcag	tggtgctacc	tccaccccag	agacgggctt	atgcccattg	27540
atctgaacat	ccagagtcac	ctttaccacg	tccctggaaa	aatacacagg	accctcaacg	27600
accgagaccg	ctggtccccg	gcctaccgcg	cgcgcaaaac	ccctaaataa	agacagcaag	27660
acacttgctt	gatcaaaaatc	caaacagagt	ctggttttta	tttatgtttt	aaaccgcatt	27720
gggaggggag	gaagccttca	gggcagaaac	ctgctggcgc	agatccaaca	gctgctgaga	27780
aacgacatta	agttccccgg	tcaaagaatc	caattgtgcc	aaaagagccg	tcaacttgct	27840
atcgcgggcg	gatgaacggg	aagctgcact	gcttgcaagc	gggctcagga	aagcaaagtc	27900
agtcacaatc	ccgcggggcg	tggtctgcagc	ggctgaagcg	gcggcgagg	ctgcagcttc	27960
caacggcggt	ccagacacgg	tctcgtaggt	caaggtagta	gagtttgcg	gcaggacggg	28020
gcgaccatca	atgctggagc	ccatcacatt	ctgacgcacc	ccggcccatg	ggggcatgcg	28080
cgttgctaaa	tatgagctca	caatgcttcc	atcaaacgag	ttggcgctca	tgggcgggcg	28140
tgctgcaaaa	cagatacaaa	actacatgag	acccccacct	tatatattct	ttcccacct	28200
taagccccgc	ccatcgatgg	caaacagcta	ttatgggtat	tatgggtgct	agcgacatga	28260
ggttgccccg	tattcagtg	cgctgatttg	tattgtctga	agttgttttt	acgttaagtt	28320
gatgcagatc	aattaatac	atacctgcgt	cataattgat	tatttgacgt	ggtttgatgg	28380
cctccacgca	cgttggtgata	tgtagatgat	aatcattatc	actttacggg	tcctttccgg	28440

```

tgatccgaca ggttacgggg cggcgacctc gcgggttttc gctatttatg aaaatatttc 28500
ggtttaaggg gtttccgttc ttcttcgtca taacttaatg tttttattta aaataccctc 28560
tgaaaagaaa ggaaacgaca ggtgctgaaa gcgaggcttt ttggcctctg tcgttttcctt 28620
tctctgtttt tgtccgtgga atgaacaatg gaagttaacg gatccaggcc gcgagcaaaa 28680
ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcggtttt ccataggctc 28740
cgccccctg acgagcatca caaaaatcaa cgctcaagtc agaggtggcg aaaccgaca 28800
ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgtc tcctgttccg 28860
accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 28920
catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 28980
gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccgtaacta tcgtcttgag 29040
tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 29100
agagcggagg atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 29160
actagaagaa cagtatttgg tatctgcgct ctgccaaagc cagttacctt cggaaaaaga 29220
gttggtagct cttgatccgg caaacaacc accgctggta gcggtggttt ttttgtttgc 29280
aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 29340
gggtctgacg ctcagtggaa cgaaaactca cgtaaaggga ttttggtcat cagattatca 29400
aaaaggatct tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt 29460
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgggc acctatctca 29520
gcgactctgc tatttcgttc atccatagtt gcctgactcc ccgtagtgtg gataactacg 29580
atacggggagg gcttaccatc cggcccaggt gctgcaatga taccgcgtga cccacgctca 29640
ccggctcctg atttatcagc aataaaccag ccagccggaa gtgccgagcg cagaagtggg 29700
cctgcaactt tatccgcctc catccagtct attagtgtt gccgggaagc tagagtaagt 29760
agttcgccag ttaatagttt tcgcaacggt gttgccattg ctacaggcat cgtggtgtca 29820
cgctcgctgt ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca 29880
tgatccccca tgttggtcaa aaaagcgggt agctccttcg gtccctccgat agttgtcaga 29940
agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact 30000
gtcatgccat ccgtaagatg cttttctgtg actggtgagt attcaaccaa gaatacggga 30060
taataccgcg ccacatagca gaactttaaa agtgctcatc attgggaaac gttcttcggg 30120
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgcg 30180
acccaagtga tcttctgcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 30240
aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 30300
tttccttttt caatattatt gaagcattta tcagggttat tgtctcatca gcggatacat 30360
atttg

```

```

<210> 5
<211> 33
<212> DNA
<213> Homo sapiens

```

```

<400> 5
gcggaattcg gcttggtgac ttagagaaca gag 33

```

```

<210> 6
<211> 33
<212> DNA
<213> Homo sapiens

```

```

<400> 6
gcgggatacct tgaaccggga ccctctcaca cta 33

```

```

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>

```

<223> derived from Adenovirus

<400> 7

actctcttcc gcatcgctgt

20

<210> 8

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 8

cttgcgactg tgactggta g

21

<210> 9

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 9

ccgcacccac tatcttcata

20

<210> 10

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> derived from Adenovirus

<400> 10

ggtgtccaaa ggttcggaga

20

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/27682

A. CLASSIFICATION OF SUBJECT MATTER IPC(7) :C12N 15/86, 15/861, 5/10, 15/11, 15/63, 15/64, 15/65; A61K 48/00 US CL :Please See Extra Sheet. According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 424/93.1, 93.2, 93.6; 435/320.1, 69.1, 455, 456, 457, 325, 369, 91.1, 91.4, 91.42 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	US 6,080,576 A (ZAMBROWICZ et al.) 27 June 2000 (27.06.00), see entire document, especially Figure 2, claims 1-3 and columns 10 and 16.	1-2, 4-11, 13
A	US 5,919,676 A (GRAHAM et al.) 06 July 1999 (06.07.99), see entire document, particularly claims 1-6 and columns 2 and 5-6.	1-34
A	AGAH et al. Gene recombination in postmitotic cells. Journal of Clinical Investigation. July 1997, Vol. 100, No. 1, pages 169-179, especially pages 171-173.	1-34
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or prior date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family	
Date of the actual completion of the international search		Date of mailing of the international search report
01 NOVEMBER 2001		31 DEC 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer David Guzo Telephone No. (703) 305-0196

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/27682

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

424/93.1, 93.2, 93.6; 435/320.1, 69.1, 455, 456, 457, 325, 369, 91.1, 91.4, 91.42

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, Dialog, NTIS, Medline, Biotech, Biosis, Biosci, Chemical Abstracts

Search terms: adenovirus, recombinase, target site, Cre, cleavage site, gene therapy

HPS Trailer Page
for
WEST

UserID: uwinkler

Printer: cm1_8e12_gblaptr

Summary

<u>Document</u>	<u>Pages</u>	<u>Printed</u>	<u>Missed</u>	<u>Copies</u>
WO000220814	117	117	0	1
Total (1)	117	117	0	-